

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ
СІКОРСЬКОГО»
ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

На правах рукопису
УДК 004.855.5:519.876.2

До захисту допущено
В. о. завідувача кафедри ММСА
_____ О. Л. Тимошук
«___» _____ 2020 р.

Магістерська дисертація

на здобуття ступеня магістра за спеціальністю 122 Комп'ютерні науки
на тему: «Моделювання і прогнозування демографічних процесів»

Виконала:
студентка II курсу, групи КА-93 мп
Гнатюк Марина Степанівна _____

Керівник: професор кафедри ММСА
д.т.н., професор Бідюк Петро Іванович _____

Рецензент: професор кафедри
інформаційної безпеки
КПІ ім. Ігоря Сікорського
д.т.н., професор Архипов О.Є. _____

Засвідчую, що в цій магістерській
дисертації немає запозичень із праць
інших авторів без відповідних посилань
Студент _____

Київ
2020

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ
СІКОРСЬКОГО»
ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ

Рівень вищої освіти — другий (магістерський)
Спеціальність — 122 «Комп'ютерні науки»

ЗАТВЕРДЖУЮ
В. о. завідувача кафедри ММСА
_____ О. Л. Тимошук
«___» _____ 2020 р.

ЗАВДАННЯ

на магістерську дисертацію студента Гнатюк Марини Степанівни

- 1. Тема дисертації:** «Моделювання і прогнозування демографічних процесів», науковий керівник дисертації Бідюк Петро Іванович д.т.н., професор, затверджені наказом по університету від 02 листопада 2020 № 3182-с
- 2. Термін подання студентом дисертації:** 14 грудня 2020 р.
- 3. Об'єкт дослідження:** демографічні процеси
- 4. Предмет дослідження:** моделі прогнозування демографічного стану населення
- 5. Перелік завдань, які потрібно розробити:**
 - 1) дослідити сучасний демографічний стан країни та провести його порівняльну характеристику.
 - 2) провести огляд сучасних підходів для побудови моделей прогнозування демографічного стану
 - 3) обрати та обґрунтувати вибір декількох моделей
 - 4) обробка вхідних даних
 - 5) реалізувати обрані моделі прогнозування
 - 6) застосувати оброблені дані на реалізованих моделях та проаналізувати отримані результати;
 - 7) розробити стартап-проект виведення на ринок результатів дослідження;
 - 8) розробити концептуальні висновки за результатами наукового дослідження.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: графіки тенденцій розвитку демографічних процесів, ілюстрації результатів прогнозування, тощо.

7. Дата видачі завдання: 1 вересня 2020 р.

Календарний план

з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації
.	Концептуальний вступ дисертації. Формулювання об'єкта, предмета, цілі, завдань, новизни, практичної значущості результатів	05.09.2020 — 13.09. 2020
.	Перший розділ. Огляд літературно- інформаційних джерел, формування нормативної бази. Характеристика об'єкта.	16.09.2020 — 27.09.2020
.	Другий розділ. Огляд моделей та методів прогнозування. Застосування математичних моделей для прогнозування демографічного стану.	30.09.2020 — 18.10.2020
.	Третій розділ. Обробка вхідних даних. Реалізація та застосування моделей. Збір та аналіз результатів.	21.10.2020 — 29.10.2020
.	Четвертий розділ. Стартап-проект	30.10.2020 — 17.11.2020
.	Концептуальні висновки. Перспективи розвитку отриманих рішень	22.11.2020 — 26.11.2020

Студент
Науковий керівник дисертації

Гнатюк М.С.
Бідюк П.І.

РЕФЕРАТ

Магістерська дисертація: 99с., 32 рис., 26 табл., 4 додатки, 26 джерел.

ДЕМОГРАФІЧНІ ПРОЦЕСИ, НЕЙРОННІ МЕРЕЖІ, НЕЛІНІЙНІ
НЕСТАЦІОНАРНІ ПРОЦЕСИ, ПРОГНОЗУВАННЯ, РЕГРЕСІЙНЕ
МОДЕЛЮВАННЯ, МГВА

Предметом вивчення протягом багатьох років є демографічний розвиток, проте навіть в час процвітання новітніх технологій, існують явища, що значно впливають на демографічні показники всього світу, і до яких людство готове.

Тому дана робота присвячена дослідженню і виділенню основних чинників, що впливають на демографічні процеси, аналізу переваг та недоліків вже існуючих інструментів та власне моделюванню демографічних процесів України.

Метою магістерської дисертації є дослідження засад прогнозування демографічних процесів, моделювання демографічних процесів. В ній наведено огляд деяких сучасних математичних методів опису динамічних процесів, а також методів апроксимації та прогнозування.

Формування нової національної стратегії, спрямованої на розвиток людського потенціалу, в поєднанні з несприятливими тенденціями демографічного відтворення населення України зумовлюють актуальність демографічних досліджень, новизною є розгляд динаміки демографічних процесів у світі з урахуванням останніх подій 2020 року, та власне оцінка їх впливу на чисельність населення.

Об'єктом дослідження є демографічні процеси, статистичні дані щодо чисельності та тенденції розвитку популяції населення України, а також виділення основних ознак, що на неї впливають.

ABSTRACT

Master's thesis consist of 99 pages, 32 images, 26 tables, 4 appendices, 26 sources

The theme: Modeling and forecasting of demographic processes.

DEMOGRAPHIC PROCESSES, NEURAL NETWORKS, NONLINEAR
NON-STATIONARY PROCESSES, FORECASTING, REGRESSION MODELING,
GMDH

The subject of study for many years is demographic development, but even during the prosperity of new technologies, there are phenomena that significantly affect the demographic indicators of the world, and for which humanity is ready.

Therefore, this work is devoted to the study and identification of the main factors influencing demographic processes, analysis of the advantages and disadvantages of existing tools and the actual modeling of demographic processes in Ukraine.

The purpose of the master's dissertation is to study the principles of forecasting demographic processes, modeling demographic processes. It provides an overview of some modern mathematical methods for describing dynamic processes, as well as methods for approximation and prediction.

Formation of a new national strategy for human development, combined with unfavorable demographic trends of population reproduction in Ukraine determine the relevance of demographic research. The novelty is the consideration of the dynamics of demographic processes in the world, taking into account the latest events of 2020, and the actual assessment of their impact on the population.

The subject of the study are demographic processes, statistical data on the number and trends of population development in different countries of the world and highlighting the main features that affect it.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	8
ВСТУП	9
РОЗДІЛ 1 ДЕМОГРАФІЧНІ ПРОЦЕСИ ЯК ОБ’ЄКТ ДОСЛІДЖЕНЬ	10
1.1 Особливості демографічного розвитку України на тлі світової тенденції	10
1.1.1 Зміна кількості населення України та її складу	10
1.1.2 Розподіл населення по території України	16
1.1.3 Міграції населення та їх вплив на демографію країни	17
1.2 Виділення основних ознак, що впливають на демографічні процеси..	18
1.3 Аналіз основних проблем демографічного розвитку України та світу	22
1.4 Вплив пандемії COVID-19 на чисельність населення в Україні.....	24
1.5 Основні напрямки та перспективи демографічної політики в Україні	27
1.5.1 Тенденції розвитку народжуваності і смертності.....	27
1.5.2 Тенденції міграції населення	29
1.5.3 Перспективи демографічного старіння	30
Висновки	31
РОЗДІЛ 2 МЕТОДИ МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ ДЕМОГРАФІЧНИХ ПРОЦЕСІВ	32
2.1 Математичні моделі для нелінійних нестационарних процесів	32
2.2 Методи апроксимації і прогнозування	37
2.2.1 Метод найменших квадратів.....	37
2.2.2 Метод групового врахування аргументів	37
2.3 Використання нейронних мереж	41
2.3.1 Класичні нейронні мережі.....	43
2.3.2 Рекурентні нейронні мережі	50

2.3.3 Нейронні мережі типу LSTM.....	53
Висновки	64
РОЗДІЛ 3 ПРОГНОЗУВАННЯ ДЕМОГРАФІЧНОГО СТАНУ В УКРАЇНІ НА ОСНОВІ ПОБУДОВАНИХ МОДЕЛЕЙ	65
3.1 Вимоги до розробленого функціоналу отримання прогнозу демографічного стану України	65
3.2 Вибір інструментальної платформи для реалізації програми	66
3.3 Розробка архітектури і функціональної схеми програми	67
3.4 Побудова математичної моделі для демографічних процесів	70
Висновки	72
РОЗДІЛ 4 СТАРТАП ПРОЕКТ.....	73
4.1 Опис ідеї проекту	73
4.2 Технологічний аудит ідеї проекту.....	75
4.3 Аналіз ринкових можливостей запуску стартап проекту	75
4.4 Розроблення ринкової стратегії проекту	81
4.5 Розроблення маркетингової програми стартап проекту	84
Висновки	87
ВИСНОВКИ.....	88
ПЕРЕЛІК ПОСИЛАНЬ.....	89
ДОДАТОК А. Динаміка чисельності населення.....	92
ДОДАТОК Б. Динаміка демографічного розвитку населення.....	94
ДОДАТОК В. Середня тривалість життя жінок і чоловіків	96
ДОДАТОК Г. Лістинг програми.....	98

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

- R^2 – коефіцієнт множинної детермінації;
- AIC – Akaike info criterion (інформаційний критерій Акайке);
- DW – Darbin-Watson (статистика Дарбіна-Уотсона);
- MAE (САП) – mean absolute error (середня абсолютна похибка);
- СААП – середня абсолютна похибка у процентах;
- СеКП – стандартне відхилення залишків, середньоквадратична похибка;
- SSE – sum of squared errors (сума квадратів похибок);
- U – коефіцієнт Тейла.
- АКФ – автокореляційна функція;
- АРКС – авторегресія з ковзним середнім;
- КС – ковзне середнє;
- СКП – сума квадратів похибок;
- МНК – метод найменших квадратів;
- ЧАКФ – часткова автокореляційна функція;
- МГВА – метод групового відсіву аргументів;
- АР(n) – авторегресія n-го порядку;

ВСТУП

Потужна, сильна та перспективна країна починається з кожного свого громадянина. Лідерські позиції на світовій арені часто визначаються економічними показниками, які у свою чергу тісно пов'язані із демографічними процесами. Демографія відіграє надзвичайно важливу роль як всередині держави, так і за її межами. Чисельність населення, тенденція його розвитку та якість стану наглядно вказує на можливі зміни напрямку вектора ведення внутрішньої політики.

Для уникнення можливих економічних, соціальних чи інших проблем та підвищення рівню життя населення, важливим є можливість оцінки тенденції розвитку демографічного стану країни, що забезпечить проведення більш глибокого аналізу її розвитку, який в результаті допоможе у прийнятті правильних рішень.

Дуже часто демографічному стану країни не надають важливого значення, ігноруються такі важливі соціальні та економічні питання, від яких залежить подальший розвиток держави в цілому.

Тому постає питання про аналіз та побудову моделей прогнозування, що дала б змогу отримати прогноз демографічних процесів для їх подальшого аналізу та використання. А також своєчасного виявлення проблем на початковому етапі, що забезпечить можливість запобігти глобальним наслідкам.

Сьогодні представлено достатньо багато програмних пакетів, які реалізують прогнозування процесів різного роду. Проте, технології розвиваються і стає можливим обробка великих масивів даних та виконання складних обчислень за достатньо задовільні часові проміжки.

РОЗДІЛ 1 ДЕМОГРАФІЧНІ ПРОЦЕСИ ЯК ОБ'ЄКТ ДОСЛІДЖЕНЬ

1.1 Особливості демографічного розвитку України на тлі світової тенденції

1.1.1 Зміна кількості населення України та її складу

Демографічний стан держави – один з найважливіших показників її рівня розвитку на міжнародній арені. Це якісна і кількісна характеристика динаміки розвитку громадян, що на пряму висвічу тенденцію економічного соціального та політичного аспектів країни.

Всі вищезгадані сфери ведення внутрішньої і зовнішньої політики взаємозв'язані і взаємозалежні, адже мають безпосередній вплив на основний ресурс кожної держави – її громадян, а ми і є основною рушійною силою, опорою і майбутнім своєї країни, її безцінне джерело праці, розвитку, відкриттів та досягнень.

Населення і є чинником, що формує трудовий потенціал, який в свою чергу – національний дохід, що дає можливість не тільки підтримувати, а й покращувати життя громадян.

Тому надзвичайно важливим є відслідковування та фіксування тенденції демографічного розвитку населення. Розробка можливості створення точного прогнозу щодо зміни її кількості, для запобігання демографічного регресу.

З початку часів, коли Україна здобула незалежність, питання щодо зміни кількості населення залишається актуальним. Вміння виділити основні ознаки та використати їх для якомога точнішого прогнозу – завдання, яке стоїть перед аналітиками, для того, щоб мати змогу вчасно прийняти правильне рішення, щодо ведення правильної внутрішньої політики.

Для України незмінним орієнтиром у питаннях ведення економічної та соціальної політики залишається Європа. Не секрет, що сучасна демографічна ситуація не є суто українським феноменом, на сьогоднішній день, тенденції, що домінують на Заході Європи притаманні більш розвинутим країнам, що здійснили демографічну революцію майже століття тому.

У сучасних громадян змінюються пріоритети ведення життя, відношення та роль сім'ї посідають не зовсім домінуюче місце у жителів нашої держави. Натомість, для українського населення важливу роль у призупиненні демографічного розвитку стали події 1990-х років, коли почалось загострення економічно та соціально-політичної кризи, 2014 року, при анексії Криму та початку війни на Сході України, також починаючи з 2019 року неабиякий вплив залишила всесвітня пандемія Covid-19, яка поставила під загрозу здоров'я та життя жителів усієї планети.

Характеризуючи сучасний демографічний розвиток країни можна виділити та охарактеризувати такі основні негативні тенденції: депопуляція в результаті зниження народжуваності та зростання смертності населення; скорочення чисельності дитячого населення, що веде за собою старіння всієї нації; загальне погіршення здоров'я населення, а звідси і надсмертність у працездатному віці; від'ємне сальдо міграцій; соціальні хвороби: ВІЛ/СНІД, Covid-19, наркоманія, алкоголізм.

Фундаментальним чинником загострення демографічної кризи в Україні є падіння народжуваності до критичного рівня. Сучасний рівень народжуваності забезпечує лише половину того, що необхідно для простого відтворення населення, тобто для заміни покоління батьків тією ж чисельністю покоління дітей. Переважно через низький рівень народжуваності Україна за 1993-2007 роки втратила більше 5 млн. населення.

Сумарний показник народжуваності зменшився з 1,9 дитини в 1989 році до 1,1 у 2001. На початок 2013 року сумарний коефіцієнт народжуваності коливається на рівні - 1,5 дитини у розрахунку на одну жінку. Дана інформація знаходиться у Додатку Б. Процес зниження народжуваності обумовлений цілою низкою причин: економічних, соціальних, психологічних, біологічних. Це є, передусім, наслідком зміни соціального статусу жінки, розширення сфери її поза сімейних інтересів, підвищення рівнів освіти та зайнятості.

Змінилося і ставлення до шлюбу як до соціального явища. Показник шлюбності з останніми роками знижується; 28,6% дітей, народжених жінками молодше 20 років, – позашлюбні; 17% сімей з дітьми є неповними. 34 З загостренням соціально-економічної кризи для багатьох українців бажання стати батьками стоїть поряд із фізичною неможливістю утримувати дитину в майбутньому [12].

Молоді пари просто елементарно не можуть дозволити собі дитину через витрати, які неодмінно виникають при народженні. З огляду на загальноєвропейські тенденції немає підстав очікувати істотного збільшення сумарних коефіцієнтів народжуваності в Україні у найближчій перспективі. У зв'язку з тим, що відкрите суспільство сприяє орієнтації на західні високі стандарти рівня життя, прагнення до них певний час буде реалізовуватись більшістю населення за рахунок відмови від народження.

На фоні загальнодержавної тенденції до зниження народжуваності неминучим стає старіння нації. Вже тепер на Заході Україна асоціюється з висловом : «Молода країна з нехарактерно старою популяцією». Частка осіб похилого віку становить в цілому по Україні понад 15% від всього населення. Так, на кожну тисячу чоловік працездатного віку припадає, в середньому, 399 непрацездатних осіб пенсійного віку.

Подальше ж старіння населення неминуче призведе до виникнення додаткових проблем в економіці та соціальній сфері по утриманню вказаної групи населення працюючими. В Україні постійно зростає рівень смертності населення. Спостерігається особливо висока смертність чоловіків у працездатному віці, яка перевищує смертність жінок в аналогічному віці майже вчетверо. Дані наведено в Додатку В.

Відбувається зростання смертності населення від інфекційних та паразитарних хвороб, реальною стала в Україні загроза епідемії туберкульозу – більш як 600 тис. її громадян хворіють сьогодні на цю хворобу. За десять місяців 2005 року кількість померлих від туберкульозу становить 9036 осіб, ВІЛ/СНІД – 2607, новоутворень (в тому числі рак) – 68686. А число смертей від хвороб системи кровообігу просто вражає – 366905. 35 Найвищі в Європі темпи поширення Covid-19 та ВІЛ/СНІДу. Особливу тривогу викликає те, що найбільшу частину інфікованих складає молодь, а серед них 15% - діти та підлітки.

Статистика кількості померлих від Covid-19 в Україні, вплив на чисельність проілюстрована на рисунку 1.1 [10].

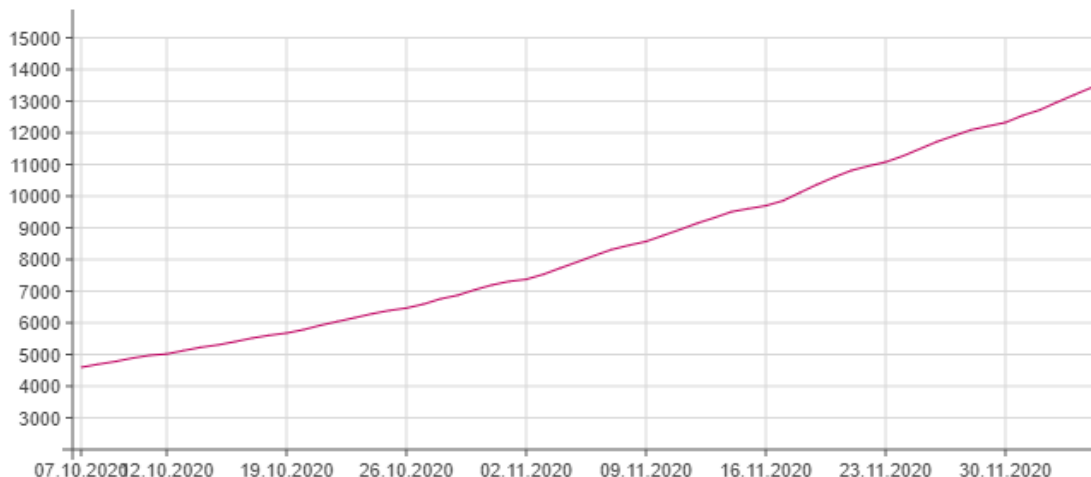


Рисунок 1.1 – Тенденція смертності від Covid-19 в Україні

Детальнішу статистику по захворюванню на Covid-19 розписано у наступних підпунктах даної роботи.

Втрати ненародженими внаслідок епідемії ВІЛ/ СНІДУ у 2015-2025 роках можуть становити 100–200 тис. осіб, що матиме істотний вплив на формування молодих поколінь. Національні та міжнародні експерти вважають, що більш як 550 тис. жителів України на сьогодні вже є ВІЛ-інфікованими [11].

Динаміка захворювань на ВІЛ/СНІД в Україні, показана на рисунку 1.2 та 1.3 відповідно.



Рисунок 1.2 – Кількість захворілих на ВІЛ по регіонах



Рисунок 1.3 – Кількість захворілих на СНІД по регіонах

Згідно прогнозів, якщо не вжити запобіжних заходів, то до 2020 року ВІЛ буде інфіковано близько 1,5 млн. українців, і його розповсюдження стане практично неконтрольованим. Зростає і масштаб соціально небезпечних хвороб - в Україні нараховується 1,2 млн. психічно хворих осіб, 720 тис. хворих на алкоголізм та 56 тис. наркоманів.

Основними причинами втрат здоров'я, зростання смертності та зниження очікуваної тривалості життя є загострення соціально-економічної кризи, низький рівень і несприятливі умови життя та праці значної частини населення, низька ефективність існуючої системи охорони здоров'я, поширеність шкідливих звичок та нехтування нормами здорового способу життя.

1.1.3 Міграції населення та їх вплив на демографію країни

Суттєво на чисельність постійного населення країни впливають міграційні процеси. Серед таких слід розрізняти насамперед два: переселення на постійне місце проживання (чи на тривалий термін) і трудову міграцію (виїзд задля заробітку).

Щодо першого процесу, то у нас ще з 1994 року зберігається від'ємне сальдо зовнішньої міграції (тобто вибуває з України більше людей, ніж прибуває на постійне місце проживання), причому до 2001 року різниця між вибулими й прибулими була досить значна і перевищувала 150 тис. осіб на рік. Потім відбувається зменшення величини модулю від'ємного сальдо, і у 2004 році перевищення було трохи більше 7 тис. осіб, а починаючи з 2005 року спостерігається міграційний приріст, який за 2005- 2013 роки становив 190,8 тис. осіб.

Кількість трудових емігрантів напряму не впливає на демографічну ситуацію (адже ці люди повертаються), однак опосередкований вплив від неї теж є. Певна частина українських заробітчани помирає за кордоном, до того ж відбувається розрив сімейних зв'язків, що веде до зростання кількості розлучень і відповідно зменшення народжуваності в Україні. Тенденція щодо міграції українців зображена на рисунку 1.5 [13].

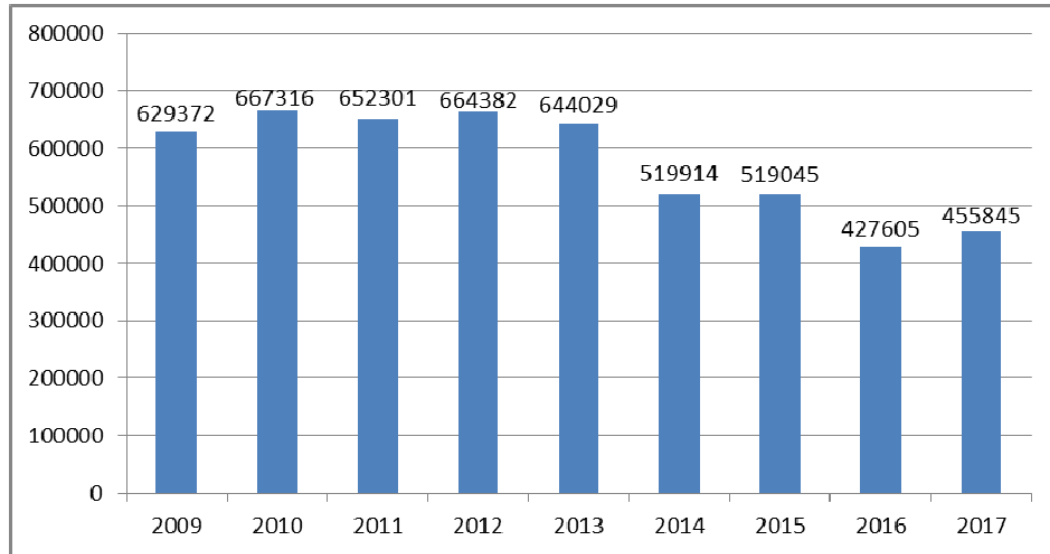


Рисунок 1.5 Міграційний рух українців

Негативним моментом є і той факт, що в еміграції беруть участь переважно молоді люди, а отже погіршується вікова структура, в Україні зменшується кількість осіб працездатного віку.

1.2 Виділення основних ознак, що впливають на демографічні процеси

Проведений аналіз сучасної демографічної ситуації, а також її динаміки протягом останніх років свідчить про наявність в Україні поряд із соціально-економічними проблемами глибокої демографічної кризи, більш інертної та практично некерованої, негативні наслідки якої для подальшого розвитку країни важко передбачити.

Статистичні порівняння з іншими країнами свідчать, що за основними показниками природних змін населення - народжуваності, смертності, природного приросту - Україна стоїть не тільки після країн Заходу, але й

республік колишнього СРСР. За рівнем народжуваності вона посідає 14 передостаннє місце серед них та країн Європи, за смертністю населення - 2-е місце серед країн континенту, за рівнем дитячої смертності (8,6 дитини на 1000 новонароджених) її показники входять у першу сумну десятку показників для європейських країн, за очікуваною тривалістю життя (68 років) - 122-е місце у світі, коли ще у 1994 р. посідала 87-е.

Загрозою для перспектив розвитку нації стають обсяги та темпи депопуляції, тобто абсолютного зменшення населення внаслідок перевищення кількості померлих над кількістю народжених. Вона почалася в Україні з 1991 р. і нині таке зменшення становить близько 1 млн. осіб. Втрачені в Україні традиції багатодітності призвели до того, що у 2013 р. сумарний показник народжуваності, тобто кількість дітей, що може народити кожна жінка упродовж усього репродуктивного періоду свого життя, становив 1,5 дитини на одну жінку (для розширеного відтворення населення необхідно мати 2,2-2,4 дитини).

Падіння життєвого рівня, що триває, невирішеність екологічних проблем, загострених наслідками Чорнобильської катастрофи, економічна криза, соціально-психологічний дискомфорт - все це примушує обмежувати розмір сім'ї. Процес зниження народжуваності характерний для багатьох розвинених країн світу, нині це загальна тенденція демографічного розвитку. Але є межа зниження народжуваності, яку не можна переступити, бо за нею процес депопуляції набуває незворотного характеру.

Україна сьогодні підійшла до цієї межі. Крім того, різко зменшилася кількість шлюбів, зростає число розлучених, удівців і особливо удовиць, збільшується частка бездітних і одно-дітних родин, а це ще більше погіршує демографічні перспективи держави. В Україні постійно зростає рівень смертності населення. Спостерігається особливо висока смертність чоловіків у працездатному віці, яка перевищує смертність жінок в аналогічному віці майже

вчетверо. Тільки за рахунок цього фактору Україна втратила у 1969-1993 рр. понад 500 тис. чоловіків у віці від 16 до 50 років. 15 Відбувається зростання смертності населення від інфекційних та паразитарних хвороб, що є ганебним явищем для цивілізованої країни.

Зростає смертність від факторів, спричинених соціальною напругою та військовим конфліктом на сході України. Інтенсифікується вимирання найбільш вразливих груп населення - дітей, жінок, осіб похилого віку. В результаті зниження народжуваності населення країни "старіє". Головною ознакою сучасної демографічної кризи є те, що в Україні відбуваються негативні зміни не тільки у кількості, але й у якості населення.

Під час обстеження стану здоров'я населення виявилось, що рівень загальної захворюваності в Україні - один із найвищих серед європейських країн. Поширюються соціальні хвороби. Так, за рівнем захворюваності алкоголізмом і наркоманією Україна перебуває на 2-му місці серед зазначеної групи країн. Існує тенденція значного погіршення здоров'я нинішніх дітей і підлітків. Лише один з чотирьох-п'яти може вважатися цілком здоровим.

Проблема поліпшення здоров'я населення сьогодні перетворилася у проблему його елементарного збереження. На погіршення основ подальшого відтворення населення України вплинули також негативні зміни в його генофонді. (Таблиця 1.1) [8].

Таблиця 1.1 - Основні демографічні показники.

	2001	2004	2007	2011	2017
Загальна чисельність наявного населення на кінець року (тис. осіб)	48457,1	47280,8	46446,1	45778,5	42760,5
Природне скорочення населення (тис. осіб)	-369,5	-334,0	-373,4	-161,9	-193
Кількість народжених (тис. осіб)	376,5	427,3	424,8	502,6	411,8
Кількість померлих (тис. осіб)	746,0	761,3	798,2	664,5	594,8
Кількість померлих дітей у віці до 1 року (осіб)	4283	4026	4523	4258	4474
Кількість шлюбів (тис.)	309,6	278,2	333,92	355,88	299,04
Кількість розлучень (тис.)	181,3	173,2	187,43	182,5	129,4
Сальдо міграції (тис. осіб)	-152,2	-7,6	9,5	17,1	9,3

Цілеспрямоване винищення найосвіченішої, професійної підготовленої частини населення України, її інтелігенції, заможних селян під час політичних репресій та голодоморів 30-40-х років, величезні втрати працездатної частини населення під час другої світової війни, організовані переселення українців до Росії, Казахстану та інших республік колишнього СРСР, зрештою, нинішній відтік інтелекту за кордон, - все це завдало і продовжує завдавати багато в чому непоправної шкоди українській нації.

1.3 Аналіз основних проблем демографічного розвитку України та світу

Демографічна проблема являє собою сукупність соціально-демографічних проблем сучасності, що зачіпають інтереси всього людства. Найважливіші проблеми народонаселення, які загрожують Україні негативними наслідками: стрімке зростання населення, або демографічний вибух, у країнах, що розвиваються, і загроза депопуляції, або демографічна криза, в економічно розвинених країнах.

До проблем народонаселення слід віднести також неконтрольовану урбанізацію в країнах, що розвиваються, кризу великих міст у деяких розвинених країнах, стихійну внутрішню й зовнішню міграцію, яка ускладнює політичні відносини між державами. Нерівномірне зростання населення в різних регіонах супроводжується інтенсивним процесом перерозподілу світового населення між ними. Частка населення економічно розвинених регіонів неухильно знижується (33,1% 1950 р., 27% 1975 р., 20% 2009 р.), тоді як частка регіонів Азії, Африки та Латинської Америки, що розвиваються, відповідно зростає.

Важливість даної проблеми полягає в тому, що такий розвиток негативно позначається на міжнародних відносинах. Подолання економічної відсталості регіонів, що розвиваються, необхідне для нормальних відносин між державами і для прогресу людства в цілому. У країнах Західної Європи, Північної Америки та Японії в 70 — на початку 90-х рр. [9].

Посилилася тенденція до різкого падіння народжуваності — значно нижче від рівня, який забезпечує просте відтворення населення. В майбутньому це загрожуватиме депопуляцією населення з її негативними соціальними наслідками. 12 Стрімке зростання населення в країнах Азії, Африки та Латинської Америки, що розвиваються, призводить до подвоєння його чисельності кожні 20-30 років і

ускладнює вирішення соціально-економічних проблем. На сьогоднішній день чисельність населення Землі складає 7,444 млрд. осіб. Протягом XX ст. кількість людей виросла у чотири рази. За оцінками ООН, щодня населення Землі збільшується на чверть мільйона, за рік ця цифра сягає 90 млн., тобто щороку чисельність збільшується на 0,7% (Рисунок 1.6) [7].



Рисунок 1.6 – Чисельність населення світу.

Такою є загальна тенденція. Ці показники зумовлені насамперед збільшенням чисельності в найбідніших регіонах світу – країнах що розвиваються, адже ледь частка зростання – 97% - припадає на Індію (1251,7 млн.), Китай (1368,7 млн.), Пакистан (199,1 млн.), Індонезію (256 млн.), Бангладеш (169 млн.), країни Африки (Нігерія – 181,6 млн.) та Латинської Америки (Бразилія – 204,3 млн.). Так, експерти вважають, що до 2050 р. В Індії мешкатиме на 100 млн. осіб більше, ніж у Китаї.

Щодо розвинених країн, то високий рівень народжуваності зберігається лише в США (321,4 млн.). 80% зростання населення в країні забезпечують

іммігранти. Згідно прогнозам, кількість людей в цій країні за наступні 50 років зросте майже на чверть.

Чисельність населення ж в Європі скорочується, незважаючи на імміграцію. Згідно з докладом Комітету по населенню Ради Європи про демографічну ситуацію, сумарне населення країн РЄ у 2010 році складало 740 млн. чол., що на 45 млн. менше, ніж у 2004 році. Зниження чисельності населення відбулося у 17 країнах, головним чином – східноєвропейських. В Литві населення скоротилося з 3,8 млн. (у 1993 році) до 3,4 млн. (у 2010 році); Білорусі, Україні, а також Греції, Італії та Швеції. Ще 5 країн мали нульовий рівень природного приросту населення. І лише дві країни (Туреччина та Ісландія) мають рівень народжуваності, що є необхідним для простого відтворення населення.

Найнижчий показник тривалості життя зафіксовано в Росії (62,5 років у чоловіків та 72 – у жінок), в Україні та Білорусі (63 — у чоловіків та 74 — у жінок). «Найстарішою» країною є Італія, «наймолодшою» — Албанія. Таким чином, на земній кулі населення розміщується вкрай нерівномірно (млн. чол.): Євразія – 4300 (Європа – 800, Азія – 3500); Африка – 800; Північна Америка – 430; Південна Америка – 350; Австралія та Океанія – 26 [10].

1.4 Вплив пандемії COVID-19 на чисельність населення в Україні.

Спалах коронавірусної інфекції COVID-19 (пневмонія нового типу) почався наприкінці грудня 2019 року. У мешканців міста Ухань провінції Хубей центрального Китаю виявили перші випадки пневмонії невідомого походження. Захворювання було пов'язане з місцевим ринком тварин та морепродуктів.

31 грудня 2019 року влада Китаю повідомила про спалах невідомої пневмонії у Всесвітню Організацію Охорони Здоров'я (ВООЗ). Уже 30 січня ВООЗ визнала спалах нового коронавірусу надзвичайною ситуацією в області громадської охорони здоров'я, що має міжнародне значення. А 11 лютого 2020 року захворювання отримало назву нової коронавірусної пневмонії — COVID-2019 (офіційна назва SARS-CoV-2).

В Україні коронавірусна інфекція Covid-19 вперше була діагностована 3 березня 2020 року в Чернівцях. 13 березня було зафіксовано перший летальний випадок внаслідок коронавірусної інфекції. На 5 грудня 2020 в Україні налічувалося 801716 занедужавших, з них померлих — 13421.

Хоч зазначається, що відсоток смертності при коронавірусній інфекції незначний, все одно в осіб, що переохворіли інфекцією, спостерігається значне погіршення роботи імунної системи, слабкість та виснаженість організму, що призводить до ризику інших захворювань, або ж повторного зараження.

Нижче наведено актуальну статистику хворих на коронавірусну інфекцію в світі на рисунках 1.7 – 1.9 відповідно.

**Поточна статистика
по коронавірусу на 5.12.2020
(у всьому світі)**

Всього інфіковано	66 380 526	
Смертельні випадки	1 527 376	2,3%
Видужали	45 962 534	69,2%
Наразі хворіють	18 890 616	28,5%

Рисунок 1.7 – Поточна статистика на 5.12.2020

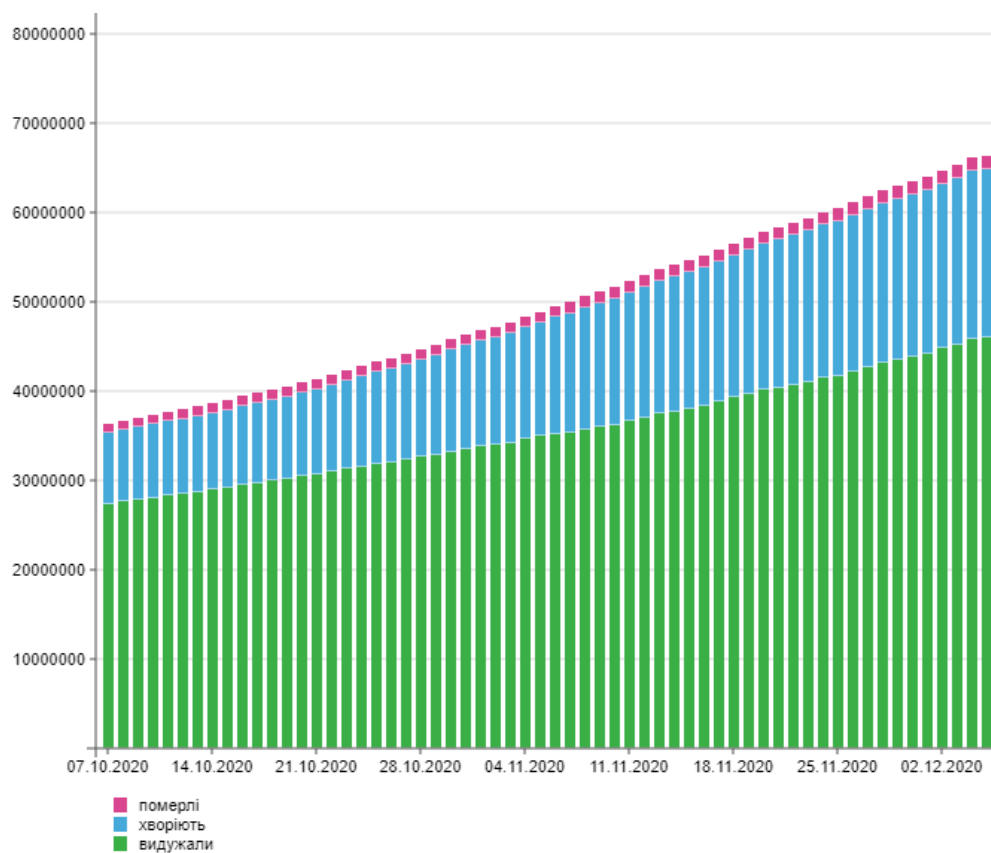


Рисунок 1.8 – Статистика хворих на Covid-19 за останні 2 місяці

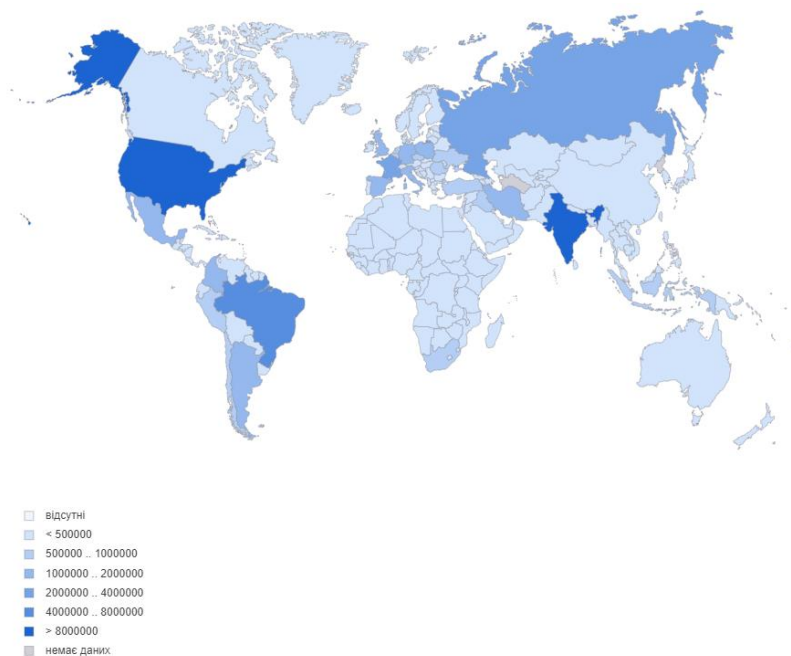


Рисунок 1.9 – Густота хворих на Covid-19 у світі

Як можемо помітити, найбільше зараз випадків захворювання у США та Індії, слід за ними – Бразилія та Росія. Кількість людей, що одужали набагато більша, ніж летальних випадків, що не може не радувати.

Щодо ситуації в Україні, то за останніми даними МОЗ, у червоній зоні знаходяться м. Київ, Одеська та Дніпропетровська області, що можемо спостерігати на рисунку 1.10.

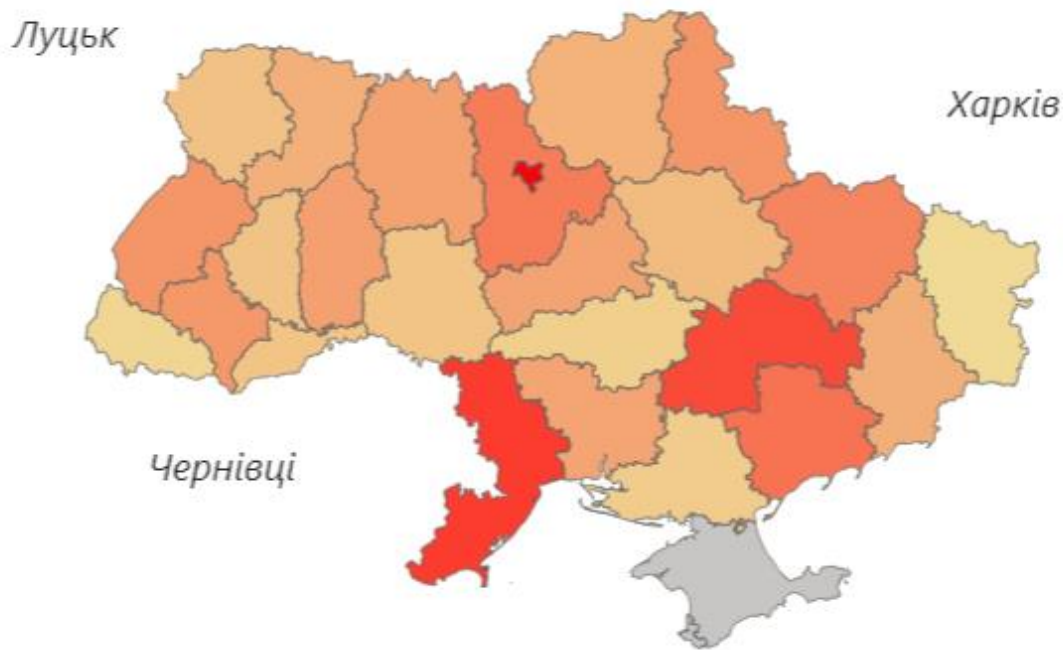


Рисунок 1.10 – Динаміка розповсюдження Covid-19 по регіонам

1.5 Основні напрямки та перспективи демографічної політики в Україні

1.5.1 Тенденції розвитку народжуваності і смертності

Загальноєвропейські тенденції народжуваності не дають підстав очікувати істотного збільшення сумарних коефіцієнтів народжуваності впродовж

наступних 10-15 років. Втрати ненародженими внаслідок епідемії ВІЛ/СНІДу протягом 2013-2020 рр. можуть становити 100-200 тис (Рисунок 1.18). У зв'язку з тим, що негативні тенденції у динаміці народжуваності посилюються, ці втрати матимуть суттєвий вплив на формування молодих поколінь. Хоча після 2015 р. спостерігається підвищення цього показника до 38 1,5 у розрахунку на одну жінку, проте це не забезпечує навіть просте заміщення поколінь [17].

Хвилеподібний характер тенденцій смертності в Україні впродовж принаймні найближчих 25 років, значні резерви зниження смертності у допрацездатних і особливо працездатних вікових групах, адаптація населення до нових соціально-економічних умов, подолання затяжної економічної кризи, успіхи медицини в лікуванні широкого кола захворювань дають підстави розраховувати на зниження ймовірності смерті: прискорене зниження - смертності немовлят (до 7,5-8%) і дітей раннього віку, вельми помірне - жінок 16-60 років (імовірність померти, не доживши до 60 років, для 16-річних дівчат скоротиться до 7-8%) і населення старше 60 років.

Найбільші зрушення (як і протягом останніх 50 років) очікуються у смертності чоловіків працездатного віку. Передбачається, що імовірність померти не доживши до 60 років для 16-річних юнаків у 2026 р. становитиме 16-21%. Практично смертність саме цього контингенту визначатиме загальний характер тенденцій смертності та тривалості життя населення України [6].

Кризовий стан дітородної активності населення даватиметься взнаки протягом життя багатьох поколінь на рівні не лише сім'ї, а й усього суспільства (навіть якщо цей процес і припиниться). Як наслідок, у 39 майбутньому спостерігатиметься хвилеподібна динаміка чисельності населення і його вікових контингентів, що може перетворитися на відчутну перепону в досягненні сталого розвитку країни.

Прогноз середньої очікуваної тривалості життя при народженні. За умови бездіяльності держави у сфері економічного забезпечення відтворення населення, відсутності належного розвитку систем охорони здоров'я, освіти, культури поширюватимуться особливо серед молоді наркоманія, алкоголізм, туберкульоз, ВІЛ/СНІД, венеричні та інші інфекційні захворювання.

Смертність впродовж певного періоду коливатиметься у межах сучасного рівня. За таких умов до 2026 р. Україна неістотно перевищить ті показники тривалості життя населення, яких вже було досягнуто у середині 80-х років [4].

1.5.2 Тенденції міграції населення

Згідно з песимістичним прогнозом за умови збереження наявних тенденцій соціально-економічного розвитку України кількість іммігрантів залишатиметься сталою, як і масштаби виїзду до інших держав. Загальне сальдо міграцій населення стабілізується на рівні близькому до сучасного.

За оптимістичним варіантом прогнозу, протягом 2015-2025 рр. можлива інтенсифікація повернення етнічних українців та раніше депортованих за національною ознакою з країн колишнього СРСР.

Щорічна кількість вибулих до економічно розвинених країн поступово зменшуватиметься, по-перше, через вичерпання етнічної складової цього потоку; по-друге, внаслідок пом'якшення дії чинників, які стимулюють від'їзд; по-третє, завдяки розширенню можливостей здійснення зворотної трудової міграції до цих країн без зміни місця проживання [5].

Підвищення рівня життя населення зробить Україну привабливою для мігрантів з різних регіонів світу. Позитивне сальдо міграцій населення постійно

зростатиме і досягне 55-60 тис. осіб у 2026 р., а коефіцієнт інтенсивності міграцій зросте з теперішніх 2,8 до 3,0-3,8%.

1.5.3 Перспективи демографічного старіння

Після відносно короткого періоду, коли 60-річну межу переходитимуть нечисленні покоління воєнних років народження, процес старіння продовжуватиметься через низьку народжуваність і зростання середньої тривалості життя, зокрема і в старших вікових групах. У селах цей процес 43 розвиватиметься повільніше, і на початок 2011 р. його показники в міських поселеннях і сільській місцевості майже зрівняються.

Після цього в селах рівень старіння знижуватиметься, а в містах - підвищуватиметься, і до 2026 р. в містах стане істотно вищим, ніж у селах. Відповідно неминуче зростатимуть рівні навантаження населення працездатного віку, що викликатиме низьку соціально-економічних проблем, які вимагатимуть посиленої уваги органів влади (Рисунок 1.11) [10].

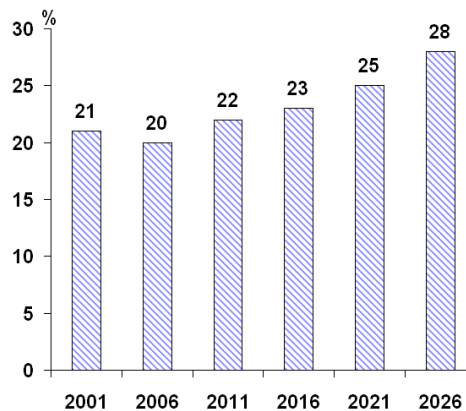


Рисунок 1.11 – Питома вага населення в віці 60 років і старше в загальній його численності, %.

Висновки

В демографічній ситуації спостерігається тенденція до зростання чисельності населення Землі. На сьогоднішній день вона складає 7,823 млрд. осіб. За оцінками ООН, щодня населення Землі збільшується на 200 000 тис. людей, за рік ця цифра сягає 90 млн., тобто щороку чисельність збільшується на 0,7%.

Але такі показники зумовлені насамперед збільшенням чисельності в найбільш розвинутих регіонах світу – країнах що розвиваються. Разом з цим, для економічно розвинутих країн характерне зниження темпів природного приросту населення. На відміну від світових тенденцій, населення України продовжує скорочуватися.

З часів останнього перепису населення 2001 року (тобто за останні 17 років), чисельність жителів країни зменшилася більш ніж на дев'ять мільйонів чоловік. Від'ємне сальдо міграції, зниження тривалості життя і стрімке падіння народжуваності, війна на сході та анексія Криму, поширення Covid-19 та ВІЛ/СНІД та соціальних хвороб стали характерними ознаками демографічних процесів останніх десятиліть і дають підстави кваліфікувати їх як ознаки демографічної кризи.

Щоб поліпшити сучасний демографічний стан, державна політика органів влади повинна насамперед поліпшувати соціально-економічні умови життя населення, лібералізувати міграційну політику та стимулювати народжуваність.

РОЗДІЛ 2 МЕТОДИ МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ ДЕМОГРАФІЧНИХ ПРОЦЕСІВ

2.1 Математичні моделі для нелінійних нестационарних процесів

Раніше вже було розглянуто основні моделі, які користуються неабиякою популярністю при прогнозуванні демографічних процесів. Нижче перераховано деякі з них, а більш детально кожна модель розглянута у [1], [2], [3] роботах.

1) Авторегресія;

$$y(k) = a_0 + a_1 y(k-1) + \dots + a_p y(k-p) = a_0 + \sum_{i=1}^p a_i y(k-i) + \varepsilon(k),$$

де p – порядок авторегресії;

k - дискретний час;

a_i – визначені на основі значень часового ряду, коефіцієнти моделі;

$\varepsilon(k)$ – випадкова величина, яка може з'являтися через похибки рівняння, вплив випадкових збурень, обчислювальні похибки тощо.

2) Парна регресія;

$$y(k) = a_0 + a_1 x(k) + \varepsilon(k),$$

де $x(k)$ – регресор

3) Множинна регресія;

$$y(k) = a_0 + a_1 x_1(k) + \dots + a_p x_p(k) + \varepsilon(k)$$

де $x_1(k), \dots, x_p(k)$ – регресори рівняння.

4) Змішана регресія;

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + b_1 x_1(k) + \dots + b_p x_p(k) + \varepsilon(k)$$

5) Авторегресія з ковзним середнім порядку (p,q);

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j \varepsilon(k-j) + \varepsilon(k)$$

6) Нелінійна регресія відносно змінних;

$$y(k) = a_0 + a_1 x(k) + \dots + a_p x^p(k) + \varepsilon(k)$$

7) Нелінійна регресія відносно параметрів;

$$y(k) = a_0 + a_1 e^{bx(k)} + \varepsilon(k)$$

8) Байєсівська мережа;

Байєсівська мережа описується у вигляді пари $\langle G, B \rangle$, де G – модифікований нециклічний граф, а B – множина параметрів, за допомогою яких визначається мережа. Параметри цієї компоненти мають наступний вигляд:

$$\Theta_{X^{(i)}|pa(X^{(i)})} = P(X^{(i)}|pa(X^{(i)})),$$

де $x^{(i)} \in X^{(i)}$ та $pa(X^{(i)}) \in Pa(X^{(i)})$, $Pa(X^{(i)})$ – набір батьків змінної $X^{(i)}$.

Для отримання повної спільної ймовірності, необхідно скористатись формулою:

$$P_B(X^{(1)}, \dots, X^{(N)}) = \prod_{i=1}^N P_B(X^{(i)}|Pa(X^{(i)})).$$

9) Авторегресія з трендовою складовою;

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^m b_j k^j + \varepsilon(k)$$

10) Авторегресія з інтегрованим ковзним середнім;

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j v(k-j) + \varepsilon(k)$$

11) Експоненційна модель;

$$y(k) = ap^k + \varepsilon(k),$$

де $y(k)$ – основна змінна процесу;

a – вирівняний тренд,

p – темп зміни рівня. Якщо $p > 1$ тренд із прискоренням, що зростає. Якщо $p < 1$, скорочення рівнів та зниження амплітуди.

Експонентний тренд практиці може розвиватися тільки на обмеженому часовому періоді

$$y(k) = \exp(\ln(a) + \ln(pk)) + \varepsilon(k).$$

12) Модель з циклічним компонентом;

$$y(k) = a_0 + \sum_{i=1}^p a_i \sin(\omega_0 i) + \sum_{j=1}^q b_j \cos(\omega_0 i) + \varepsilon(k),$$

де $y_n(k)$ – основна (залежна) змінна;

$a_0, \{a_i, b_i\}, i = 1, \dots, n$ – коефіцієнти моделі;

Можна використовувати будь-яку детерміновану функцію, для опису тренду, головне наявність її відповідності характеру його зміни у часі.

13) Поліноміальна модель;

$$y(k) = a_0 + a_1 k + a_2 k^2 + \dots + a_m k^m + \varepsilon(k),$$

де $y(k)$ – основна (залежна) змінна процесу;

$a_i, i = 0, \dots, m$ – коефіцієнти моделі;

$\varepsilon(k)$ – випадковий процес, через вплив зовнішніх збурень, де час є тим визначальним фактором, що впливає на незалежну змінну у та входить до моделі у вигляді регресорів з різними степенями.

14) Моделі гетероскедастичних процесів.

Гетероскедастичними називаються моделі, дисперсія яких є змінна з часом. Рівняння для умовної дисперсії таких моделей має наступний вигляд:

$$h(k) = \beta_0 + \varepsilon^2(k-1) + \varepsilon_1(k),$$

де $\varepsilon^2(k)$ – квадрат залишків;

$h(k)$ – умовна дисперсія процесу;

$\varepsilon_1(k)$ – похибка моделі;

k – дискретний час.

Узагальнена авторегресійна умовно гетероскедастична модель:

$$h(k) = \beta_0 + \sum_{i=1}^p \beta_i \varepsilon^2(k-i) + \sum_{i=1}^q a_i h(k-i) + \varepsilon_1(k),$$

де $\alpha, \beta > 0$ для уникнення від’ємних значень умовних дисперсій.

Авторегресійна умовно гетероскедастична модель порядку p

$$h(k) = \beta_0 + \sum_{i=1}^p \beta_i \varepsilon^2(k-i) + \varepsilon_1(k).$$

Якщо серед коренів характеристичного рівняння для обраної авторегресійної частини є одиночний, то дані формули описують тренд і, накладені на нього коливання.

Крім засобів знаходження умовних ймовірностей в орієнтованих графах існує чимало інших байєсівських методологій, таких як моделі динамічних процесів, моделі із симетричними зв’язками, моделі із прихованими змінними, що детальніше описані у роботі [13].

2.2 Методи апроксимації і прогнозування

2.2.1 Метод найменших квадратів

Метод найменших квадратів використовується для пошуку найближчого розв'язку системи. Суть даного методу полягає в знаходженні коефіцієнтів лінійної залежності, при яких функція, що описує суму квадратів відхилень від прямої, що розглядається, приймає найменше значення.

Для цього достатньо знайти екстремум функції змінних. Складається система рівнянь з однаковою кількістю рівнянь та невідомих, після чого знаходимо похідні по змінним і прирівнюємо їх до нуля.

Після вирішення системи. Ми отримуємо формули для знаходження коефіцієнтів за допомогою методу найменших квадратів [26].

$$\hat{\theta} = [X^T X]^{-1} X^T y,$$

де $\theta[p]$ – вектор оцінок параметрів вимірності p ;

$X[N \times p]$ – матриця вимірів;

$y[N]$ – вектор вимірів залежної змінної. Елементи матриці вимірів обчислюються по-своєму для кожної конкретної моделі.

2.2.2 Метод групового врахування аргументів

Розв'язок практичних задач та розробка теоретичних питань Метода Групового Урахування Аргументів (МГУА) призвели до появи широкого спектру

обчислювальних алгоритмів, кожен з яких використовується за певних умов застосування. Вибір алгоритму залежить як від точності інформації, яка представлена у вибірці експериментальних даних, так і від виду задачі, яка розв'язується.

Метод групового урахування аргументів базується на завданні правил ускладнення моделі, системі опорних функцій, критерію селекції та методу регуляризації згідно зовнішнім критеріям.

Розглянемо формальну постановку задачі:

Задана множина даних: вхідні змінні $\{X(1), X(2), \dots, X(M)\}$ і вихідна змінна $\{y(1), y(2), \dots, y(M)\}$, де $X = \{x_1, x_2, \dots, x_n\}$ – n -вимірний вектор, M – число точок спостереження.

Необхідно на основі спостережних даних побудувати модель

$Y = F(x_1, x_2, \dots, x_n)$, яка буде найменш складна. Відмінними особливостями даної задачі є:

- вид функціональної залежності невідомий, а визначений лише клас моделей, наприклад, поліном довільного порядку або гармонічний ряд;
- коротка вибірка даних;
- часові ряди $x_i(t)$ в загальному випадку нестационарні.

В такому випадку застосування класичних методів статистичного аналізу, наприклад, регресійного чи дисперсійного аналізу, неможливо, і необхідно використовувати нестандартні методи, наприклад, основані на застосуванні ідей штучного інтелекту.

Метод бере ідеї з біології, а саме механізми еволюції:

- схрещення або гібридизація предків і генерація потомків;
- селекція та відбір кращих.

Майже усі відомі алгоритми МГУА використовують поліноміальні опорні функції. Загальний зв'язок між вхідними та вихідними змінними знаходиться у

вигляді функціонального ряду Вольтера, дискретний аналог якого відомий як поліном Колмогорова-Габора.

$$Y = a_0 + \sum_{i=1}^N a_i x_i + \sum_{i=1}^N \sum_{j=1}^N a_{ij} x_i x_j + \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N a_{ijk} x_i x_j x_k + \dots$$

Оскільки з допомогою такого поліному можна досягти досить точної апроксимації будь-якої диференційованої функції.

Ця складна залежність замінюється множиною простих функцій:

$y_1 = f(x_1, x_2); y_2 = f(x_1, x_3); \dots y_s = f(x_{N-1}, x_N)$, де $s = C^2$, причому функція f усюди однакова.

Отже, отримавши перший ряд селекції, нові змінні подаються тепер як вхідні параметри, і отримуємо другий ряд селекції:

$$z_1 = f(y_1, y_2); z_2 = f(y_1, y_3); \dots z_p = f(y_{s-1}, y_s), \text{ де } p = C_s^2$$

В даній роботі розглядаються три види часткових описів:

- лінійний: $f(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j$
- неповний квадратичний: $f(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j$,
- квадратичний: $f(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2$,

При побудові моделі (при знаходженні значень коефіцієнтів) в якості критерію використовується критерій регулярності:

$$\bar{\varepsilon}^2 = \frac{1}{N} \sum_{i=1}^N (y_i - f(x_i))^2$$

Необхідно забезпечити $\varepsilon^2 \rightarrow \min$. Розглянемо покроковий алгоритм МГУА.

- 1) Для кожної пари x_i та x_j будуються часткові описи:

$$\hat{Y}^{(s)} = \varphi(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2, \text{ де } s=1 \dots C_N^2.$$

- 2) Обчислюються коефіцієнти цих моделей по МНК, використовуючи навчальну вибірку. Тобто, знаходимо $\hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{a}_5$.

- 3) Потім на перевіірочній вибірці для кожної з цих моделей шукаємо

$$\bar{\varepsilon}_s^2 = \frac{1}{N_{test}} \sum_{k=1}^{N_{test}} [Y(k) - \hat{Y}_k^{(s)}]$$

де $Y(k)$ – дійсне значення вихідної змінної в k -тій точці перевіірочної вибірки; $\hat{Y}^{(s)}$ – вихідне значення в k -тій точці перевіірочної вибірки у відповідності з s -тою моделлю.

Відбираємо F кращих моделей, і подаємо на другий ряд. Шукаємо

$$\hat{Z}^{(P)} = \varphi^{(2)}(x_i, x_j) = a_0^{(2)} + a_1^{(2)} y_i + a_2^{(2)} y_j + a_3^{(2)} y_i y_j + a_4^{(2)} y_i^2 + a_5^{(2)} y_j^2$$

Оцінка тут така ж, як на першому ряді.

Процес конструювання рядів повторюється до тих пір, поки середній квадрат помилки буде падати. Коли на шарі m отримаємо збільшення помилки ε^2 , то процес закінчуємо [24].

2.3 Використання нейронних мереж

Штучні нейронні мережі є електронними моделями нейронної структури мозку, який головним чином навчається з досвіду. Природній аналог доводить, що множина проблем, які поки що не підвладні розв'язуванню наявними комп'ютерами, можуть бути ефективно вирішені блоками нейромереж.

Базовим елементом мозку людини є специфічні клітини, відомі як нейрони, що здатні запам'ятовувати, думати і застосовувати попередній досвід до кожної дії, що докорінно відрізняє їх від решти клітин організму.

Сила людського розуму залежить від числа базових компонент, різноманіття з'єднань між ними, а також від генетичного програмування і навчання. Індивідуальний нейрон є складним, має свої складові, підсистеми та механізми керування і передає інформацію через велику кількість електрохімічних зв'язків.

Налічують біля сотні різних класів нейронів. Разом нейрони та з'єднання між ними формують недвійковий, нестійкий та несинхронний процес, що відрізняється від процесу обчислень традиційних комп'ютерів. Штучні нейромережі моделюють лише найголовніші елементи складного мозку, що надихає науковців та розробників до нових шляхів розв'язування проблеми.

Розглянемо біологічний нейрон (рисунок 2.1) [26].

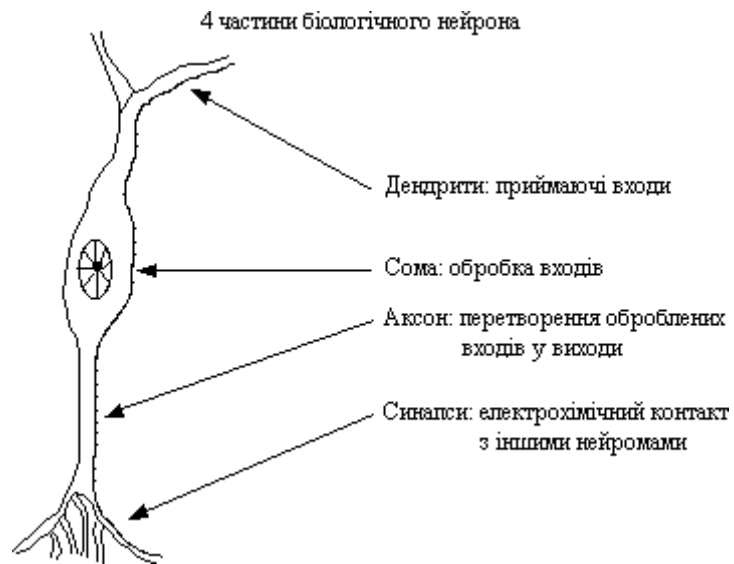


Рисунок 2.1 – Схема біологічного нейрона.

Нейрон (нервова клітина) складається з тіла клітини – соми, і двох типів зовнішніх деревоподібних відгалужень: аксона і дендритів. Тіло клітини вміщує ядро, що містить інформацію про властивості нейрону, і плазму, яка продукує необхідні для нейрону матеріали. Нейрон отримує сигнали (імпульси) від інших нейронів через дендрити (приймачі) і передає сигнали, згенеровані тілом клітки, вздовж аксона (передавач), що наприкінці розгалужується на волокна. На закінченнях волокон знаходяться синапси.

Синапс є функціональним вузлом між двома нейронами (волокно аксона одного нейрона і дендрит іншого). Коли імпульс досягає синаптичного закінчення, продукуються хімічні речовини, названі нейротрансмітерами. Нейротрансмітери проходять через синаптичну щілину, збуджуючи або загальмовуючи, у залежності від типу синапсу, здатність нейрона-приймача генерувати електричні імпульси. Результативність синапсу налаштовується минаючими через нього сигналами, тому синапси навчаються в залежності від активності процесів, у яких вони приймають участь. Нейрони взаємодіють за

допомогою короткої серії імпульсів. Повідомлення передається за допомогою частотно-імпульсної модуляції.

Останні експериментальні дослідження доводять, що біологічні нейрони структурно складніші, ніж спрощене пояснення, наведене вище і значно складніші, ніж існуючі штучні нейрони, які є елементами сучасних штучних нейронних мереж. Оскільки нейрофізіологія надає науковцям розширене розуміння дії нейронів, а технологія обчислень постійно вдосконалюється, розробники мереж мають необмежений простір для вдосконалення моделей біологічного мозку.

2.3.1 Класичні нейронні мережі

Розглянемо штучний нейрон. Базовий модуль нейронних мереж штучний нейрон моделює чотири основні функції природного нейрона (рис. 2.2).

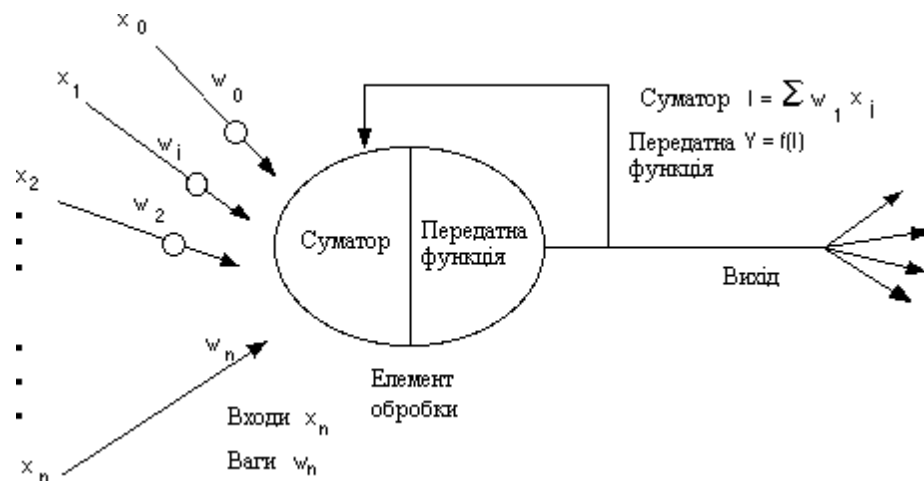


Рисунок 2.2 – Базовий штучний нейрон.

Вхідні сигнали x_n зважені ваговими коефіцієнтами з'єднання w_n додаються, проходять через передатну функцію, генерують результат і виводяться.

У наявних на цей час пакетах програм штучні нейрони називаються "елементами обробки" і мають набагато більше можливостей, ніж простий штучний нейрон, описаний вище. На рис. 2.3 зображена детальна схема спрощеного штучного нейрону.

Модифіковані входи передаються на функцію сумування, яка переважно тільки сумує добутки. Проте можна обрати багато різних операцій, такі як середнє, найбільше, найменше, OR, AND, тощо, які могли б виробляти деяку кількість різних значень.

Окрім того, більшість комерційних програм дозволяють інженерам-програмістам створювати власні функції сумування за допомогою підпрограм, закодованих на мові високого рівня). Інколи функція сумування ускладнюється додаванням функції активації, яка дозволяє функції сумування оперувати в часі.

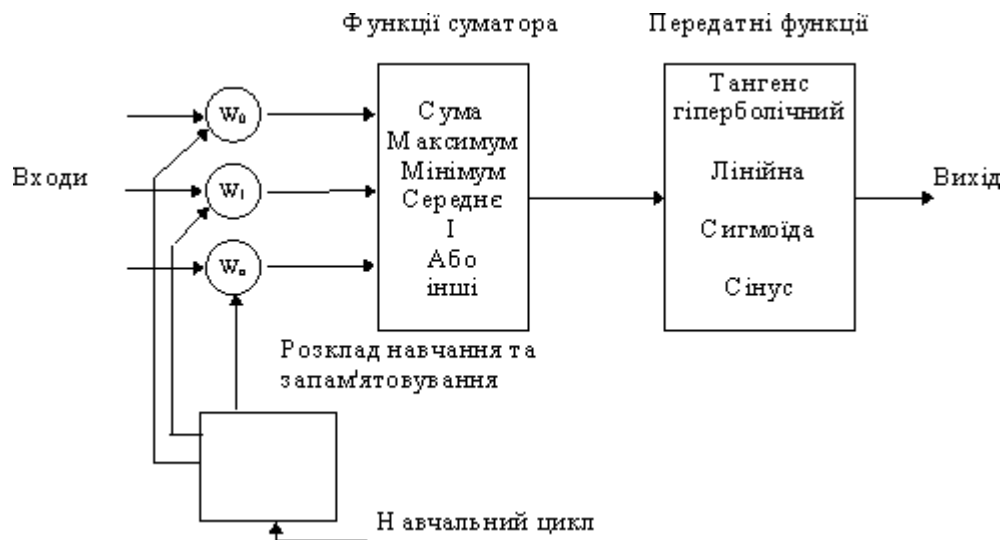


Рисунок 2.3 – Модель "елементу обробки".

В будь-якому з цих випадків, вихід функції сумування надсилається у передатну функцію і скеровує весь ряд на дійсний вихід (0 або 1, -1 або 1, або яке-небудь інше число) за допомогою певного алгоритму. В існуючих нейромережах в якості передатних функцій можуть бути використані сигмоїда, синус, гіперболічний тангенс та ін. Приклад того, як працює передаточна функція показаний на рис. 2.4.

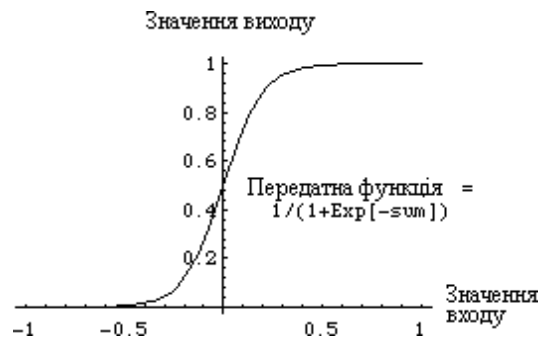


Рисунок 2.4 – Сигмоїдна передаточна функція.

Після обробки сигналу, нейрон на виході має результат передатної функції, який надходить на входи інших нейронів або до зовнішнього з'єднання, як це передбачається структурою нейромережі [23].

Всі штучні нейромережі конструюються з базового формуючого блоку штучного нейрону. Існуючі різноманітності і фундаментальні відмінності, є підставою мистецтва талановитих розробників для реалізації ефективних нейромереж.

Інша частина створення і використання нейромереж стосується нескінченної кількості зв'язків, що пов'язують окремі нейрони. Групування у мозку людини відбувається так, що інформація обробляється динамічним, інтерактивним та самоорганізуючим шляхом. Біологічні нейронні мережі створені у тривимірному просторі з мікроскопічних компонент і здатні до різноманітних з'єднань. Але для

створеної людиною мережі існують фізичні обмеження. Існуючі на даний час нейромережі є групуванням штучних нейронів. Це групування обумовлено створенням з'єднаних між собою прошарків.

На рис. 2.5 показана типова структура штучних нейромереж.

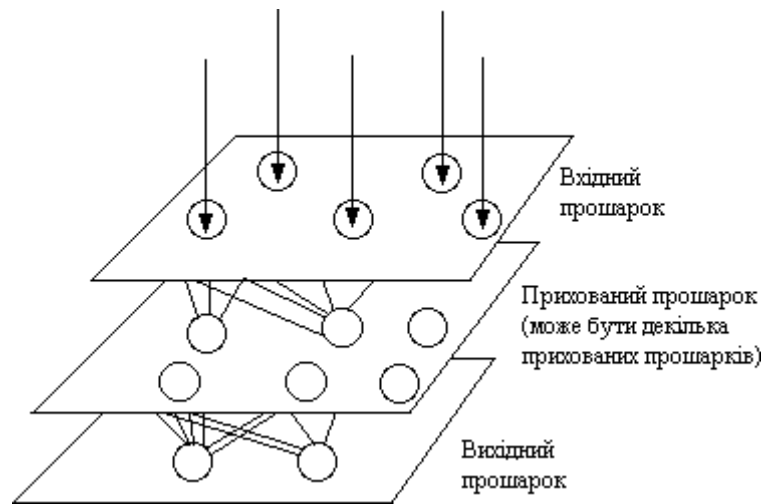


Рисунок 2.5 – Діаграма простої нейронної мережі.

Хоча існують мережі, які містять лише один прошарок, або навіть один елемент, більшість застосувань вимагають мережі, які містять як мінімум три нормальних типи прошарків – вхідний, прихований та вихідний. Прошарок вхідних нейронів отримує дані або з вхідних файлів, або безпосередньо з електронних давачів. Вихідний прошарок пересилає інформацію безпосередньо до зовнішнього середовища, до вторинного комп'ютерного процесу, або до інших пристроїв.

Між цими двома прошарками може бути багато прихованих прошарків, які містять багато нейронів у різноманітних зв'язаних структурах. Входи та виходи кожного з прихованих нейронів просто йдуть до інших нейронів.

Напрямок зв'язку від одного нейрону до іншого є важливим аспектом нейромереж. У більшості мереж кожен нейрон прихованого прошарку отримує

сигнали від всіх нейронів попереднього прошарку та звичайно від нейронів вхідного прошарку. Після виконання операцій над сигналами, нейрон передає свій вихід до всіх нейронів наступних прошарків, забезпечуючи шлях передачі вперед (feedforward) на вихід. При зворотному зв'язку, вихід нейронів прошарку скеровується до нейронів попереднього прошарку (рис. 2.6).

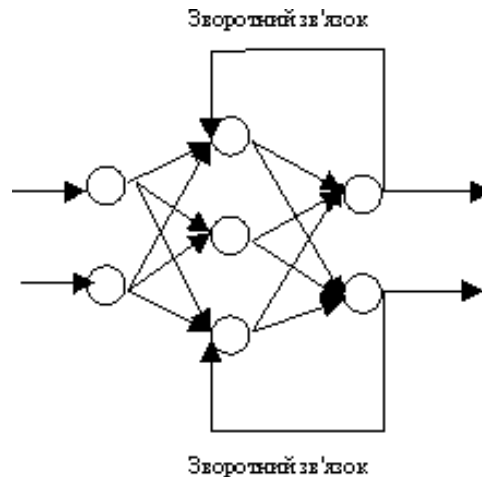


Рисунок 2.6 – Мережа зі зворотнім зв'язком

Шлях, яким нейрони з'єднуються між собою має значний вплив на роботу мережі. Більшість пакетів професійної розробки програмного забезпечення дозволяють користувачу додавати, вилучати та керувати з'єднаннями як завгодно. Постійно корегуючи параметри, зв'язки можна робити як збуджуючими так і гальмуючими.

Отже, нейронні мережі - це системи штучного інтелекту, здатні до самонавчання в процесі вирішення різних задач.

Розглянемо особливості навчання штучної нейронної мережі. Здатність до навчання є фундаментальною властивістю мозку. Процес навчання може розглядатися як визначення архітектури мережі і налаштування ваг зв'язків для ефективного виконання спеціальної задачі.

Нейромережа налаштовує ваги зв'язків по наявній навчальній множині. Властивість мережі навчатися на прикладах робить їх більш привабливими в порівнянні із системами, які функціонують згідно визначеній системі правил, сформульованої експертами.

Існують три загальні парадигми навчання: "з вчителем", "без вчителя" (самонавчання) і змішана. У першому випадку нейромережа має у своєму розпорядженні правильні відповіді (виходи мережі) на кожен вхідний приклад. Ваги налаштовуються так, щоб мережа виробляла відповіді як можна більш близькі до відомих правильних відповідей.

Навчання без вчителя не вимагає знання правильних відповідей на кожен приклад навчальної вибірки. У цьому випадку розкривається внутрішня структура даних та кореляція між зразками в навчальній множині, що дозволяє розподілити зразки по категоріях. При змішаному навчанні частина ваг визначається за допомогою навчання зі вчителем, у той час як інша визначається за допомогою самонавчання.

Модель будь-нейронної мережі складається, таким чином:

- з елементів, що обробляють сигнали (нейронів);
- зв'язків між нейронами;
- навчальних схем або правил.

Кожен нейрон мережі володіє власною вагою (позитивною, негативною) і так званої функцією активації. Нейрон обробляє вхідні сигнали, підсумовуючи входи з вагами і трансформуючи отриманий результат у вихідний сигнал за допомогою функції активації (лінійної, нелінійної, стохастичної). Нейрони мережі взаємодіють один з одним в залежності від топології зв'язків. Останні бувають повними (кожен з кожним), частковими (протилежність повної зв'язку), з петлями зворотних зв'язків, без таких. У кожній задачі потрібен свій вибір архітектури мережі.

Наявність зворотного зв'язку впливає на здатність до навчання мережі (покращуючи її), тоді як ступінь внутрішніх зв'язків визначає паралелізм обчислень. Коли нейромережі стали популярні, вони складалися з одного або двох шарів нейронів - вхідного і / або вихідного. Діапазон завдань, що вирішується такими мережами, був обмежений. Сучасні багат шарові нейромережі потужніші, зокрема, через те, що вони формують внутрішнє представлення задачі в так званих прихованих шарах.

Найбільш популярний алгоритм навчання нейромережі - це зворотне розповсюдження помилки, що складається з взаємозв'язаних процесів. У прямому процесі вхідний сигнал проходить через мережу, генеруючи певний вихід. У зворотному процесі помилка (різниця між бажаним і отриманим виходом) передається від вихідних шарів до вхідних з одночасною модифікацією зв'язків нейронів так, щоб (при подальшому прогоні інформації через мережу) помилка на вихідному шарі зменшилася. У цьому полягає навчання мережі.

Незважаючи на переваги нейронних мереж в часткових галузях над традиційними обчисленнями, існуючі нейромережі є не досконалими рішеннями. Вони навчаються і можуть робити "помилки". Окрім того, не можна гарантувати, що розроблена мережа є оптимальною мережею. Застосування нейромереж вимагає від розробника виконання ряду умов.

Ці умови включають:

- множину даних, що включає інформацію, яка може характеризувати проблему;
- відповідно встановлену за розміром множину даних для навчання і тестування мережі;
- розуміння базової природи проблеми, яка буде вирішена;
- вибір функції суматора, передатної функції та методів навчання;
- розуміння інструментальних засобів розробника;

– відповідна потужність обробки.

Новий шлях обчислень вимагає вмінь розробника поза межами традиційних обчислень. Спочатку, обчислення були лише апаратними і інженери робили його працюючим. Потім, були спеціалісти з програмного забезпечення: програмісти, системні інженери, спеціалісти по базах даних та проектувальники. Тепер є нейронні архітектори.

Новий професіонал повинен мати кваліфікацію відмінну від його попередників. Наприклад, він повинен знати статистику для вибору і оцінювання навчальних і тестових множин. Логічне мислення сучасних інженерів програмного забезпечення, їх емпіричне вміння та інтуїтивне відчуття гарантує створення ефективних нейромереж [24].

2.3.2 Рекурентні нейронні мережі

Найбільш сучасні (і вважаються найбільш «вдалими») з нейронних мереж беруть початок зі структури, званої багат шаровим перцептроном (математичної моделі мозку – штучні нейронні мережі прямого поширення з проміжними шарами). При цьому з часів своєї появи вони зазнали значних змін - і штучні нейронні мережі «нового покоління» влаштовані набагато простіше своїх попередників, при тому, що вони дозволяють з успіхом вирішувати завдання запам'ятовування послідовностей.

Так, наприклад, популярна на сьогоднішній день мережа Елмана влаштована таким чином, що зворотний сигнал з внутрішнього шару надходить не на «головні» вхідні нейрони, а на додаткові входи - так званий контекст. Ці нейрони зберігають інформацію про попередній вхідний вектор (стимул);

виходить, що вихідний сигнал (реакція мережі) залежить не тільки від поточного стимулу, а й від попереднього (рис. 2.7).

На практиці тимчасові ряди мають сильну зашумленість - що, природно, викликає проблеми при спробі прогнозування. Знизити ступінь помилки дозволяє використання колекцій мереж прямого поширення - однак це суттєво збільшує не тільки складність самої структури, але і час її навчання.

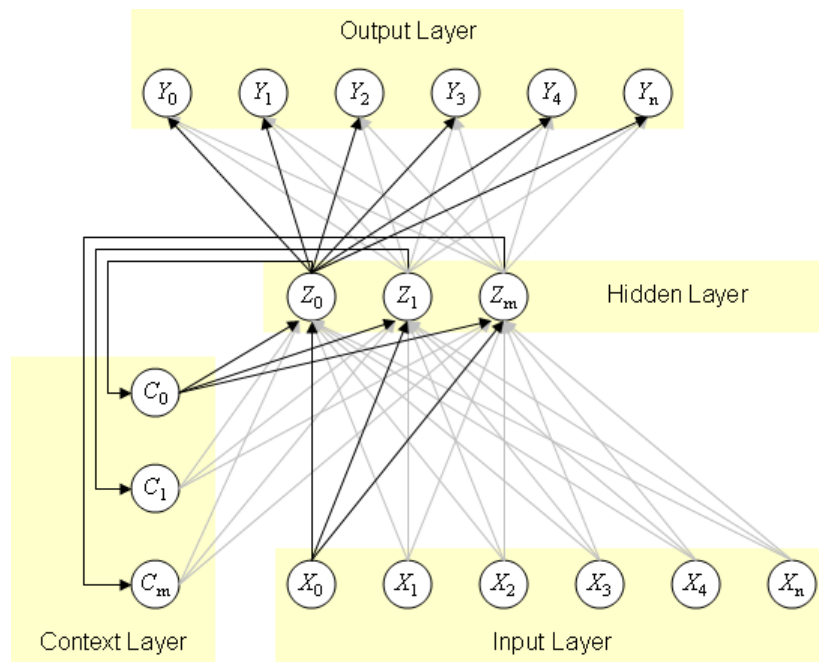


Рисунок 2.7 – Мережа Елмана.

Метод з використання рекурентних мереж успішно використовується для прогнозування фінансово економічних процесів, оскільки саме рекурентна мережа вивчає закономірності в послідовності величин, що необхідно для роботи з тимчасовими рядами.

Недоліком такої мережі є тривалий час навчання. Пошарові мережі, на кожен ідентичний вхідний патерн відповідають одним і тим же патерном на

виході. Рекурентні мережі можуть відповідати на один і той же вхідний патерн по-різному, в залежності від того, який патерн був попереднім.

Таким чином, на них суттєво впливає послідовність навчальних прикладів. Іншими словами, рекурентні мережі можуть бути навчені так само, як стандартні мережі зі зворотним розповсюдженням помилки, проте навчальні приклади повинні бути впорядковані і не можуть пропонуватися мережі в випадково вибраному порядку.

Істотна різниця з пошаровими мережами полягає в наявності у мережі Елмана - Джордана додаткового блоку, що зберігає інформацію про попередні входи. Він може бути інтерпретований як блок довготривалої пам'яті нейромережі.

Нейрони прихованого шару мережі Елмана - Джордана мають логістичну функцію активації:

$$f(x) = \frac{1}{1 + e^{-x}}$$

тоді як нейрони вихідного шару мають лінійною функцією активації. Така комбінація дозволяє апроксимувати будь-яку функцію з кінцевим числом розривів із заданою точністю.

По суті, єдиною вимогою до мережі є достатня кількість нейронів в прихованому шарі, оскільки їх число повинне зростати пропорційно складності аналізованих даних.

Найбільш важким у використанні нейромереж є вибір моменту зупинки навчання. Якщо мережу навчати недовго, то вона не вивчить закономірності у вибірці навчальних прикладів.

Якщо мережу навчати занадто довго, то вона вивчить приклади з шумами з як завгодно високою точністю, але виявиться нездатною узагальнювати

приклад (тобто буде діяти схожим чином на даних, що не входили в навчальну множину).

Використання рекурентної мережі Елмана дозволяє вирішувати задачу прогнозування навіть на сильно зашумлених тимчасових рядах (це особливо важливо для бізнесу). У загальному випадку ця штучна нейронна мережа є структурою з трьох шарів, а також набору додаткових «контекстних» елементів (входів). Зворотні зв'язки йдуть від прихованого шару до цих елементів; кожна зв'язок має фіксований вагу, що дорівнює одиниці.

На кожному часовому відрізку входні дані розподіляються по нейронам в прямому напрямку; потім на них застосовується навчальне правило. Завдяки фіксованим зворотним зв'язкам, контекстні елементи завжди зберігають копію значень з прихованого шару за попередній крок (оскільки вони відправляються в зворотному напрямку ще до застосування навчального правила).

Таким чином, шум тимчасового ряду поступово нівелюється, і разом з ним мінімізується і помилка: ми отримуємо прогноз, який в загальному випадку буде точніше, ніж результат класичного підходу, що західні роботи підтверджують експериментально [21].

2.3.3 Нейронні мережі типу LSTM

Одна з ідей, яка робить рекурентні нейронні мережі настільки привабливими, полягає в тому, що вони могли б використовувати отриману в минулому інформацію для поточних завдань. Наприклад, вони могли б використовувати попередні кадри відео для розуміння наступних. Якби РНС

могли це робити, вони були б у край корисні. Але чи можуть вони? Залежить від ситуації.

Іноді нам досить недавньої інформації, щоб виконувати поточну задачу. Наприклад, уявімо модель мови, яка намагається передбачити наступне слово, ґрунтуючись на попередніх. Якщо ми намагаємося передбачити останнє слово в реченні "Хмари на небі", нам не потрібен більше ніякого контексту - досить очевидно, що в кінці речення мова йде про небо.

У таких випадках, де невеликий проміжок між необхідною інформацією і місцем, де вона потрібна, рекурентні нейронні мережі можуть навчитися використовувати інформацію, отриману раніше (рис. 2.8).

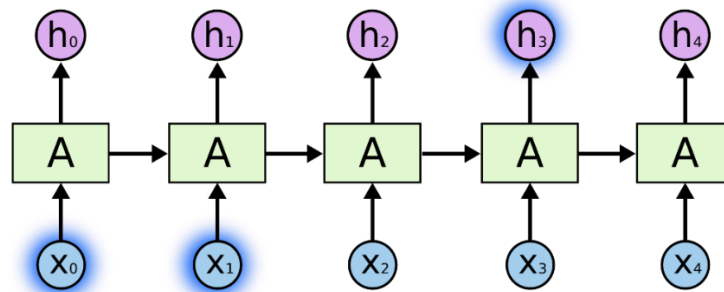


Рисунок 2.8 – Загальна схема рекурентної нейронної мережі.

Але також бувають випадки, коли нам потрібен більш широкий контекст. Припустимо, потрібно передбачити останнє слово в тексті "Я виріс у Франції ... Я вільно розмовляю французькою". Недавня інформація підказує, що наступне слово, ймовірно, назва мови, але якщо ми хочемо уточнити, якого саме, нам потрібен попередній контекст аж до інформації про Францію. Зовсім рідко проміжок між необхідною інформацією і місцем, де вона потрібна, стає дуже великим.

На жаль, при зростанні проміжку, рекурентні мережі стають нездатні навчитися поєднувати інформацію (рис. 2.9).

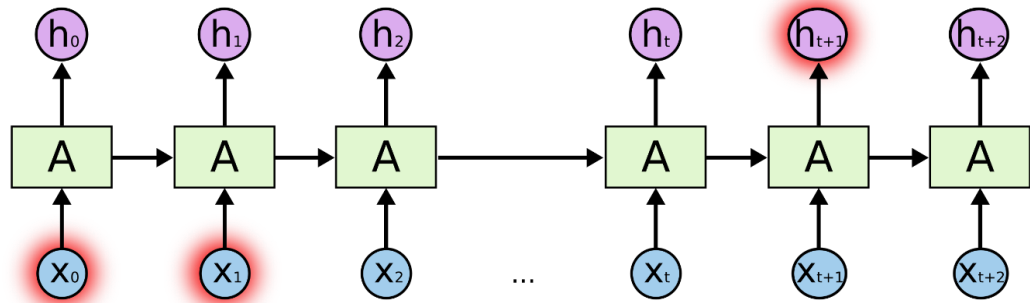


Рисунок 2.9 – Рекурентні мережі з великими проміжками.

Теоретично, рекурентні нейронні мережі здатні обробляти такі довготривалі залежності. Людина може ретельно підібрати їх параметри, щоб вирішувати іграшкові проблеми такої форми. На жаль, на практиці, не схоже, щоб рекурентні нейронні мережі були здатні розв'язати таку проблему. Проблема була глибоко вивчена Хохрейтером (Hochreiter) (1991) і Бенджі і ін. (1994). Їм вдалося знайти деякі досить фундаментальні причини, чому це може бути важко. На щастя, нейронні мережі типу LSTM не мають цієї проблеми.

Мережі довго-короткостроковій пам'яті (Long Short Term Memory) - зазвичай просто називають "LSTM" - особливий вид РНМ, здатних до навчання довгостроковим залежностям.

Вони були запропоновані Хохрейтером і Шмідхубер (Schmidhuber) (1997) і доопрацьовані і популяризовані іншими в подальшій роботі. Вони працюють неймовірно добре на великій різноманітності проблем і в даний момент широко застосовуються.

LSTM спеціально спроектовані таким чином, щоб уникнути проблеми довгострокових залежностей. Запам'ятовувати інформацію на тривалий період

часу - це практично їх поведінка за замовчуванням, а не щось таке, що вони тільки намагаються зробити.

Все рекурентні нейронні мережі мають форму ланцюга повторюваних модулів (repeating module) нейронної мережі. У стандартній РНМ ці повторювані модулі будуть мати дуже просту структуру, наприклад, всього один шар гіперболічного тангенса (\tanh) (рис. 2.10).

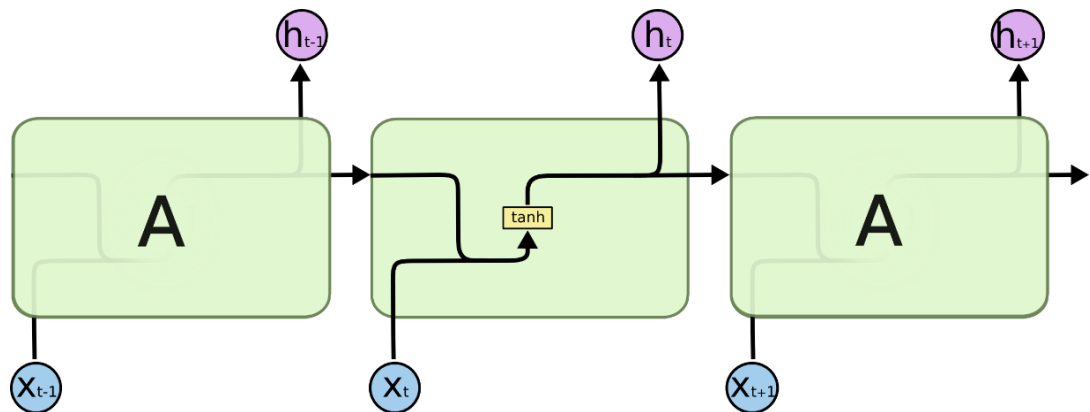


Рисунок 2.10 – Структура повторюваного модуля РНМ.

LSTM теж мають таку ланцюгову структуру, але повторюваний модуль має іншу побудову. Замість одного нейронного шару їх чотири, причому вони взаємодіють особливим чином (рис. 2.11).

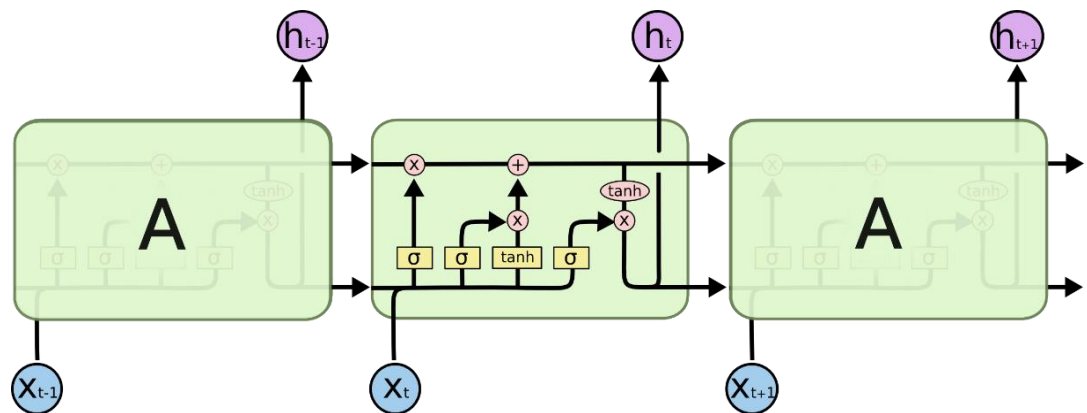


Рисунок 2.11 – Структура повторюваного модуля LSTM.

У діаграмі вище кожна лінія передає цілий вектор від виходу одного вузла до входів інших. Рожеві кола представляють поточкові оператори, такі як додавання векторів, в той час, як жовті прямокутники - це навчені шари нейронної мережі. Лінії, що зливаються позначають конкатенацію, в той час як розгалужені лінії позначають, що їх вміст копіюється, і копії відправляються в різні місця. Ключ до LSTM - клітинний стан (cell state) - горизонтальна лінія, що проходить крізь верхню частину діаграми.

Клітинний стан - це щось типу стрічки конвеєра. Вона рухається прямо вздовж всього ланцюга тільки з невеликими лінійними взаємодіями. Інформація може просто текти по ній без змін (рис. 2.12).

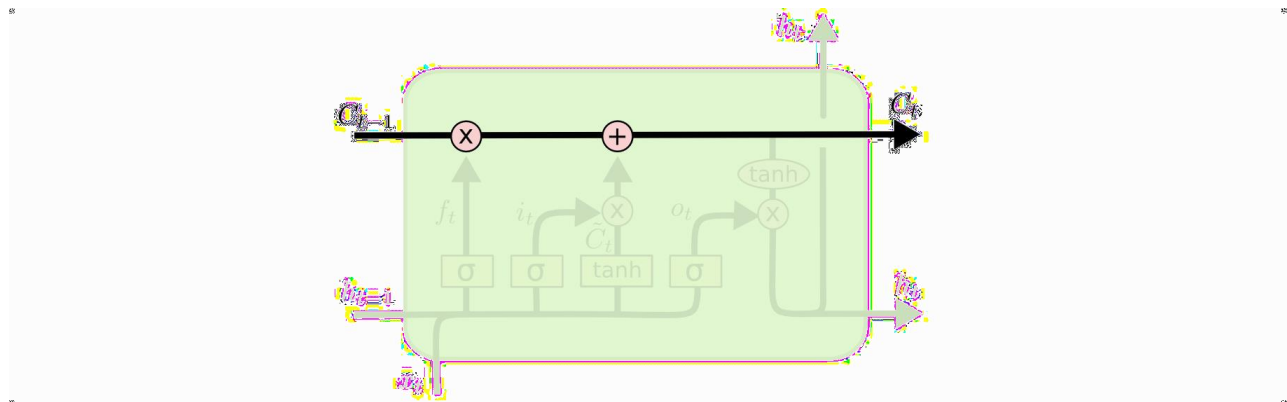


Рисунок 2.12 – Клітинний стан LSTM.

LSTM має здатність видаляти або додавати інформацію до клітинного стану, проте ця здатність ретельно регулюється структурами, що називаються вентилями (gates).

Вентилі - це спосіб вибірково пропускати інформацію. Вони складені з сигмоїдного шару нейромережі і операції поточкового множення (pointwise multiplication) (рис. 2.13).

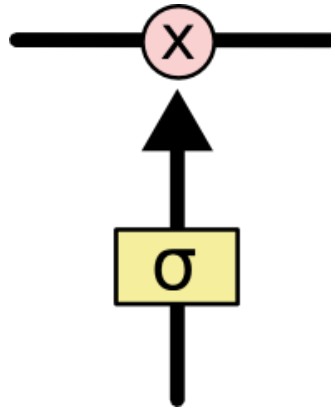


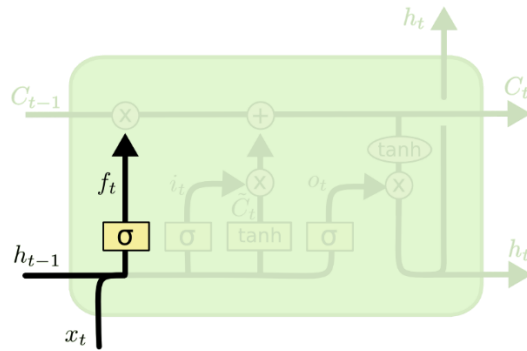
Рисунок 2.13 – Приклад вентиля LSTM.

Сигмоїдний шар подає на вихід числа між нулем і одиницею, описуючи таким чином, наскільки кожен компонент повинен бути пропущений крізь вентиль. Нуль - "нічого не пропускати", один - "пропускати все".

LSTM має три таких вентиля, щоб захищати і контролювати клітинний стан.

Першим кроком в LSTM треба вирішити яку інформацію ми збираємося викинути з клітинного стану. Це рішення приймається сигмоїдним шаром, що називається "забуваючим вентилям" ("forget gate layer"). Він дивиться на $h(t-1)$ і $x(t)$ і подає на вихід число між 0 і 1 для кожного числа в клітинному стані $C(t-1)$. Одиниця означає "збережи це повністю", в той час як нуль означає "позбудься цього повністю" (рис. 2.14).

Повертаючись до прикладу мовної моделі, яка намагається передбачити наступне слово, ґрунтуючись на всіх попередніх, бачимо що у такій проблемі клітинний стан може включати рід підмета, що дозволить використовувати правильні форми займенників. Коли ми бачимо новий підмет, ми забуваємо рід попереднього підмета.

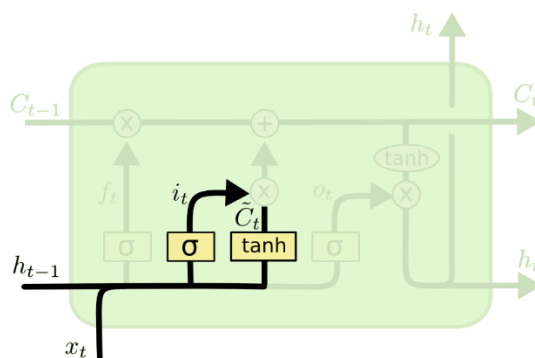


$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Рисунок 2.14 – Забуваючий вентиль LSTM.

Наступним кроком треба вирішити, яку нову інформацію ми збираємося зберегти в клітинному стані. Цей крок складається з двох частин. По-перше, сигмоїдний шар, що називається "вхідним вентиляем" ("input gate layer"), вирішує, які значення ми оновимо. Далі, шар гірпоболічного тангенса створює вектор кандидатів на нові значення $C(t)$, який може бути доданий до стану. На наступному кроці ми з'єднаємо ці дві частини, щоб створити оновлення для стану (рис. 2.15).

У прикладі з нашої мовної моделлю ми б хотіли додати рід нового підмета до клітинного стану, щоб замінити рід старого, який ми повинні забути.



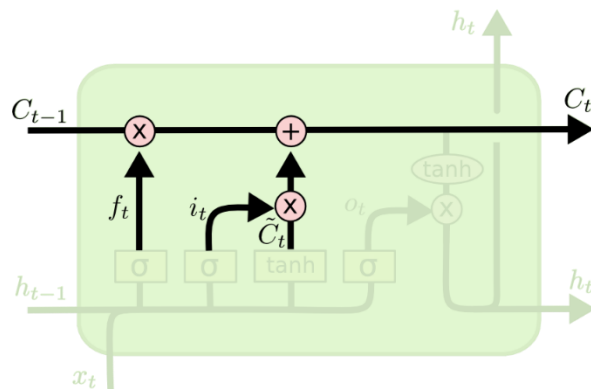
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Рисунок 2.15 – Вхідний вентиль LSTM.

Далі необхідно оновити старий клітинний стан, $C(t-1)$, новим клітинним станом $C(t)$. Всі рішення вже прийняті на попередніх кроках, залишилося тільки

зробити це. Ми множимо старий стан на $f(t)$, забуваючи все, що ми раніше вирішили забути. Далі ми додаємо $i(t) * C(t)$. Це нові кандидати в значення, масштабовані відповідно до того, як сильно ми вирішили оновити кожне значення стану (рис. 2.16).



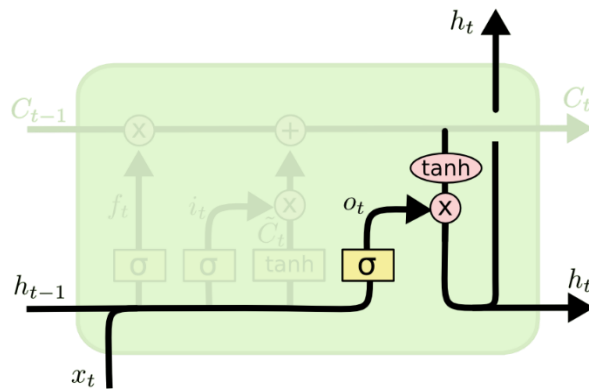
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Рисунок 2.16 – Оновлення клітинного стану LSTM.

У випадку з мовної моделлю, це якраз те місце, де ми втрачаємо інформацію про рід старого підмета і додаємо нову інформацію, як вирішили на попередніх кроках.

Нарешті, нам потрібно вирішити, який результат ми збираємося подати на вихід. Цей результат буде заснований на нашому клітинному стані, але буде його відфільтрованою версією. Спочатку ми запускаємо сигмоїдний шар, який вирішує, які частини клітинного стану ми збираємося відправити на вихід.

Потім ми пропускаємо клітинний стан крізь гіперболічний тангенс (\tanh) (щоб вмістити значення в проміжок від -1 до 1) і множимо його на вихід сигмоїдного вентиля, так що ми відправляємо на вихід тільки ті частини, які ми хочемо (рис. 2.17).



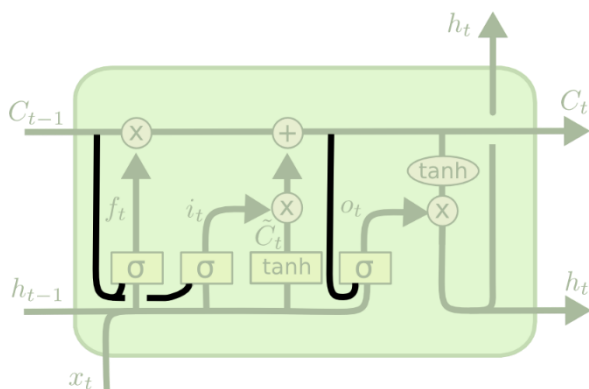
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

Рисунок 2.17 – Формування виходу LSTM шару.

У прикладі з мовною моделлю, якщо вона тільки що бачила підмет, вона могла б подати на вихід інформацію, що відноситься до дієслова (в разі, якщо таке слово саме дієслово). Наприклад, вона, можливо, подасть на вихід число підмета (єдине або множинне). Таким чином, ми будемо знати, яка форма дієслова повинна бути підставлена (якщо звичайно далі йде саме дієслово).

Один з популярних варіантів LSTM, запропонований Герсом і Шмідгубером (Gers & Schmidhuber) у 2000 додає "оглядові з'єднання" ("peerhole connections") (рис. 2.18). Це означає, що ми дозволяємо вентилям "підглядати" за клітинним станом.



$$f_t = \sigma(W_f \cdot [C_{t-1}, h_{t-1}, x_t] + b_f)$$

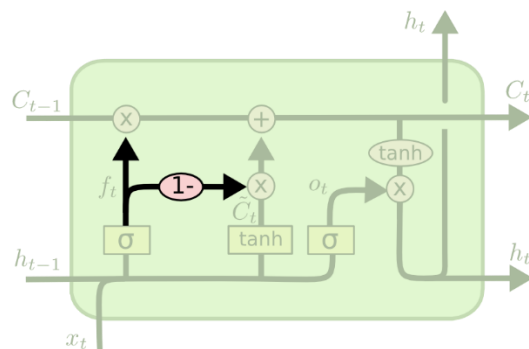
$$i_t = \sigma(W_i \cdot [C_{t-1}, h_{t-1}, x_t] + b_i)$$

$$o_t = \sigma(W_o \cdot [C_t, h_{t-1}, x_t] + b_o)$$

Рисунок 2.18 – «Оглядові з'єднання» LSTM шару.

Діаграма вище додає очі (оглядові отвори) для всіх вентилів, але в багатьох публікаціях ви зустрінете лише деякі з них, а не всі відразу.

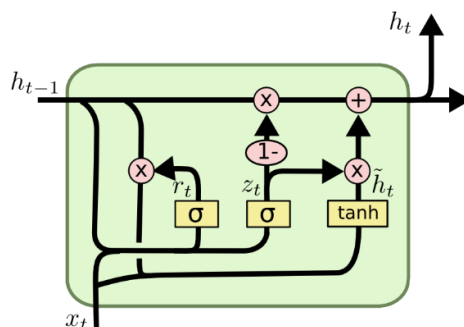
Інша варіація - використання спарених забуваючих і входних вентилів (рис. 2.19). Замість того, щоб незалежно вирішувати що забути і куди ми повинні додати нову інформацію, ми приймаємо ці рішення одночасно. Ми забуваємо щось тільки в тому випадку, коли ми отримуємо щось інше на це місце. Ми отримуємо на вхід нові значення тільки коли забуваємо щось старе.



$$C_t = f_t * C_{t-1} + (1 - f_t) * \tilde{C}_t$$

Рисунок 2.19 – Спарені забуваючі і входні вентиляі.

Дещо більш істотно відрізняється від LSTM вентиляна рекурентна одиниця (Gated Recurrent Unit) або GRU, запропонована Чо (Cho) та ін в 2014 (рис 2.20). Вона поєднує забуваючі і входні вентиляі в один "відновлювальний ventиль" ("update gate"). Вона також зливає клітинний стан з прихованим шаром і вносить деякі інші зміни. Модель, що виходить в результаті, простіше, ніж звичайна модель LSTM і вона набирає популярність.



$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Рисунок 2.20 – Gated Recurrent Unit.

Це тільки деякі з найбільш популярних варіантів LSTM. Є безліч інших, наприклад глибинно-вентильні PHM (Depth Gated RNNs) Яо (Yao) і ін. (2015).

Існує і зовсім інший підхід до вивчення довготривалих залежностей, наприклад годинні PHM (Clockwork RNNs) Коутніка (Koutnik) і ін. (2014 року).

Який з цих варіантів найкращий? Чи мають ці відмінності якесь значення? Греффі (Greff) і ін. (2015) зробили чудове порівняння популярних варіантів і виявили, що вони все приблизно однакові. Джозефовіч (Jozefowicz) і ін. (2015) протестували понад тисячу архітектур PHM і виявили, що деякі з них працюють краще, ніж LSTM на певних завданнях.

Висновки

В даному розділі було наведено деякі сучасні моделі і методи прогнозування, що використовуються при прогнозуванні демографічних процесів. Згадано моделі, що використовувались для побудови СППР для прогнозування нелінійних гетероскедастичних процесів. Дані підходи стали особливо популярними за останні декілька років.

Також було описано методи апроксимації та прогнозування, розглянуто новий підхід, а саме використання нейронних мереж. Описано їх основні види і суть їхньої роботи.

Серед представлених моделей для створення чіткого прогнозу демографічних процесів, коротко описано декілька з них, а саме: моделі авторегресії, авторегресії з ковзним середнім, парної регресії, змішаної регресії, множинної регресії, регресії з нелінійними змінними, регресії з нелінійними параметрами, моделі для гетероскедастичних процесів, байєсівські мережі, експоненційні моделі, моделі з циклічним компонентом та авторегресії із трендовою складовою.

Також було описано метод групового врахування аргументів та розглянуто класичні нейронні мережі, рекурентні нейронні мережі та нейронні мережі типу LSTM.

РОЗДІЛ 3 ПРОГНОЗУВАННЯ ДЕМОГРАФІЧНОГО СТАНУ В УКРАЇНІ НА ОСНОВІ ПОБУДОВАНИХ МОДЕЛЕЙ

3.1 Вимоги до розробленого функціоналу отримання прогнозу демографічного стану України

Для того, щоб отримати прогноз високої якості, необхідно дотримуватись технологій, що ґрунтуються на застосуванні динаміки методів структуризації задач, методів обробки статистичних даних, моделей, процесів та критеріїв їх оцінки.

Застосування такого підходу гарантує отримання на виході адекватний прогноз та прийняті рішення, що на ньому ґрунтуються.

Проте, для забезпечення обробки великої кількості інформації, необхідно мати зручний, потужний та новітній інструментарій. Розглянемо декілька основних вимог користувача, що необхідні для створення СППР:

1) гнучка та зрозуміла мова програмування, яка забезпечить оптимальний розподіл ресурсів та пам'яті, необхідних для виконання процесу роботи, що забезпечить мінімізацію часу обробки запитів користувача;

2) максимально простий та зрозумілий інтерфейс користувача, для введення попередніх даних, демонстрації процесу обробки даних, представлення кінцевих результатів, тощо;

3) можливість збереження отриманих результатів.

Система підтримки прийняття рішень була розроблена мовою програмування python, містить в собі програмні модулі, що відповідають за зчитування та обробку вхідних даних, розбиття їх на тренувальну та основну вибірки, проведення тренування, власне прогнозування, обчислення характеристик якості моделі та прогнозу, вивід результатів, а також можливість

їх збереження та представлення часу, для обчислення ефективності роботи програми.

Для можливості використання розробленої програми необхідно бути забезпеченим наступними характеристиками ПК:

1. Частота процесора 1.6 -3.4 ГГц.
2. Об'єм оперативної пам'яті 8 ГБ і більше.
3. Інтегрований графічний адаптер.
4. Операційна система: Windows 10, Linux.
5. Роздільна здатність екрану: (1920x1080) Full HD.
6. Розширення, що дають можливість переглядати файли формату *.csv, *.xlsx.

3.2 Вибір інструментальної платформи для реалізації програми

СППР реалізована мовою програмування python, яка є досить популярною сьогодні. Python нова мова програмування, але не зважаючи на це, вона входить у трійку найпопулярніших мов програмування. До переваг python належать: читабельність, простота та прямота коду, легка інтерпретація, динамічна типізація.

Легка система імпортування, вже вбудовані інструменти та величезна база бібліотек спрощують роботу, і дають можливість сфокусуватись на побудові алгоритму для вирішення конкретних задач.

За допомогою інтерпретатора швидкість проведення експериментів зростає вразі. Мова програмування Python надає можливість написання веб додатків, десктоп та мобільних додатків. Тобто це мова широкого призначення.

3.3 Розробка архітектури і функціональної схеми програми

Для проектування архітектури СППР важливим фактором є вибір підходу. Існують різні види підходів:

- інформаційний;

Передбачує належність СППР до класу інформаційних систем, створених для поліпшення характеру діяльності управління, шляхом застосування засобів інформаційних технологій. Основними компонентами таких систем є інтерфейс «користувач» - «система», база даних, а також база моделей.

- підхід заснований на знаннях;

Полягає в об'єднанні технологій підтримки рішень та штучного інтелекту. Основна відмінність даного підходу полягає у спроможності системи зрозуміти проблему, зреагувати на запит користувача, обробити його та видати результат; Структура даної СППР вміщує в собі наступні ключові елементи: мовна система, система знань, система обробки.

- інструментальний підхід;

Залежно від обраного напрямку, розглядаються наступні три рівні системи підтримки прийняття рішень: спеціалізовані СППР, генератори СППР, інструментарій СППР. Управління інтерфейсом такої СППР має реалізовуватись за допомогою меню, мови команд та мови запитань-відповідей. Аналіз даних контролюється в базі моделей, а управління представленнями підтримує різноманітні образи користувача стосовно проблеми, яку необхідно вирішити.

Після аналізу вищевказаних пунктів, за основу було обрано інформаційний підхід, оскільки він найбільш зрозумілий для користувача, і не використовує

додаткових ресурсів та пам'яті, які можна використати власне при обробці статистичних даних та реалізації прогнозу.

Після вибору підходу проектування постає питання про вибір типу архітектури СППР. Розрізняють наступні типи архітектур:

- текстово-орієнтовані;

Застарілий вид архітектури, що містить мовну систему, базу знань, систему обробки даних, генерування результатів та їх представлення. Властивістю такої архітектури є гіпертекстова підтримка, що встановлює зв'язок між знаннями, які містяться у текстових файлах.

- орієнтовані на бази даних;

Архітектура СППР орієнтована на роботу з реляційними базами даних, що забезпечує обробку строго структурованих знань.

- орієнтовані на використання електронних таблиць;

Даний вид архітектури представляє основу процедурних знань, що служать командами для системи обробки даних та генерування результатів. Тип архітектури призначений для виконання аналізу «що, якщо», для порівняння результатів дій над даними в таблицях. Іншими словами така архітектура прискорює порівняння всіх альтернатив.

- основані на правилах;

Особливість полягає у збереженні правил та опису станів до бази знань даних, та можливості формування висновку на основі вказаних правил. Окрім стандартних запитів, архітектура СППР дозволяє на основі правил створювати запити стосовно прийняття рішень (поради та пояснення деяких фактів).

- гібридні;

Очевидно, що така архітектура є універсальнішою та дає набагато більше можливостей для процесу обробки інформації та видачі результатів. Гібридна архітектура є гнучкою, що спрощує інтеграцію СППР в майбутньому [20].

Саме гібридний тип архітектури був обраний для створення СППР, і він має наступний узагальнений вигляд, зображений на рисунку 3.1

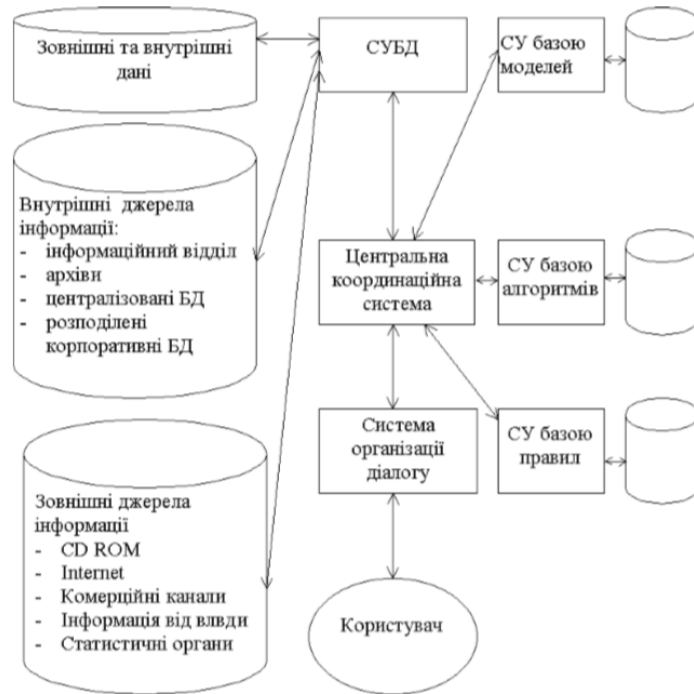


Рисунок 3.1 – Узагальнена архітектура гібридної СППР

Ще однією складовою при створенні СППР є функціональна схема програми. У розробленій СППР присутні наступні модулі, що реалізують всі необхідні функції для коректної роботи програми:

- Config.py – модуль, що реалізовує зручну роботу із файлами.
- Metric.py – модуль з метрикою, що використовується при обчисленнях.
- Preprocess.py – модуль, що готує необхідні дані для прогнозу.
- Train_test.py – модуль, де формуються тестова вибірка і йде процес навчання моделі.
- Time.py – модуль, призначений для визначення часу роботи програми, за для аналізу ефективності вцілому.
- Predict.py – модуль, у якому власне вибирається країна та запускається

процес прогнозування, виведення та запису результатів.

Також, у ході роботи було використано наступні бібліотеки:

- а). numpy;
- б). pandas;
- в). xlrd;
- г). lightgbm;
- д). scikit-learn;
- е). tqdm.

На рисунку 3.2 зображено дерево структури проекту, з ієрархією модулів та програмних папок.

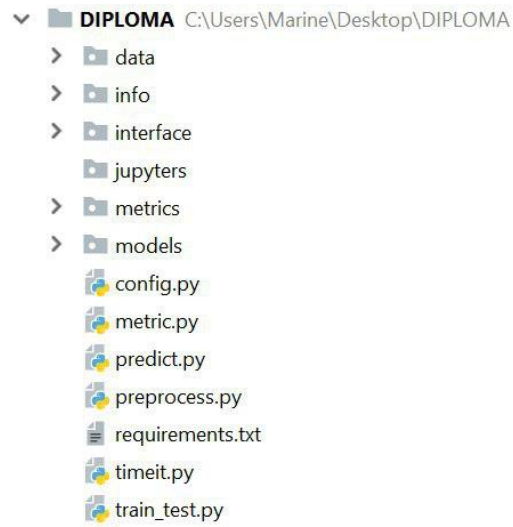


Рисунок 3.2 – Дерево структури проекту

3.4 Побудова математичної моделі для демографічних процесів

У даній роботі було використано модель MLPRegressor. Ця модель оптимізує квадратні втрати за допомогою LBFGS або стохастичного градієнтного

спуску. На основі даних зібраних з різних джерел статистичних даних, було здійснено прогноз динаміки народонаселення. При цьому були застосовані різні підходи до моделювання динамічних процесів.

Було використано нейронну мережу прямого розповсюдження з 6-ма внутрішніми шарами по 125 нейронів кожен. Для навчання моделі використовувався метод Adam. Оскільки вважається найшвидшим відомим на даний момент. Як активаційна функція використовувалась функція ReLU через найбільшу коректність даних та простоти у розумінні і використанні.

Використовувалось три бібліотеки: pandas, numpy і scikit-learn (в коді прописується як sklearn). Перша використовувалася для зчитування даних з файла. Друга та третя використовувалися для додаткової обробки та підготовки даних. Третя бібліотека має широкі можливості як до підготовки даних, так і для побудови різних прогностичних та класифікаційних моделей, зокрема, на основі нейронної мережі.

Висновки

Після дослідження, збору та обробки статистичних даних, аналізу якості обраних моделей прогнозування та розробки коду, було створено СППР, що дає адекватний прогноз чисельності населення обраної країни світу на N років вперед. СППР була розроблена на основі інформаційного підходу, з використанням гібридного типу архітектури, мовою програмування python.

Було обрано за основу модель MLPRegressor. Ця модель оптимізує квадратні втрати за допомогою LBFGS або стохастичного градієнтного спуску. На основі даних зібраних з різних джерел статистичних даних, було здійснено прогноз динаміки народонаселення. При цьому були застосовані різні підходи до моделювання динамічних процесів.

Розроблено максимально простий та зрозумілий інтерфейс користувача, за допомогою якого робота з СППР стала більш приємною.

РОЗДІЛ 4 СТАРТАП ПРОЕКТ

4.1 Опис ідеї проекту

На сьогодні великої популярності набуває такий вид підприємництва як стартап. Стартап-проект — є комерційним проектом, який знаходиться в стані розробки, або нещодавно вийшов на ринок. Характерною особливістю стартапу, що відрізняє його від малого бізнесу, є оригінальність та інновації, він не може бути копією вже реалізованих ідей.

При цьому проект не обов'язково повинен бути масштабного характеру, головне, щоби він був креативним, а його завдання — спрощувати людям будь-які дії в їх повсякденному житті. Наразі, з появою Інтернету та сучасних технологій, стало простіше заходити на ринок, знаходити інвесторів та споживачів. З'явилося набагато більше можливостей для розвитку свого проекту за кордоном, ніж раніше.

Проте розробка стартапу є досить ризикованим завданням. Не всім вдається довести свій стартап-проект до ринкового впровадження. За статистикою успіху досягає лише 10–20 % від усіх стартап-проектів.

Запуск стартапу передбачає цілий ряд обов'язкових дій, у межах яких визначають ринкові перспективи стартапу, графік розробки, принципи організації виробництва, заходи з залучення інвесторів та аналіз ризиків [18].

Ідея та реалізація стартап проекту зображена у табл. 4.1 - 4.22.

Таблиця 4.1 – Опис ідеї стартап проекту

Зміст ідеї	Напрямки застосування	Вимоги користувача
Система прийняття рішень в сфері демографічних процесів	Надання допомоги ОПР у прийнятті рішень в сфері демографічних процесів	Отримання допомоги у прийнятті рішень в сфері демографічних процесів

Таблиця 4.2 – Визначення сильних, слабких та нейтральних характеристик ідеї проекту

№ п/п	Техніко-економічні характеристики ідеї	(потенційні) товари/концепції конкурентів			W (слабка сторона)	N (нейтральна сторона)	S (сильна сторона)
		Мій проект	СППР SAS	СППР Oracle			
1.	Ціна	Низька	Висока	Висока			+
2.	Ефективність	Висока	Висока	Висока		+	
3.	Функціонал	Вузький	Широкий	Широкий	+		

4.2 Технологічний аудит ідеї проекту

Таблиця 4.3 – Технологічна здійсненність ідеї проекту

№ п/п	Ідея проекту	Технології її реалізації	Наявність технологій	Доступність технологій
1.	Розробка ПЗ, а саме СППР в сфері демографічних процесів	C++	Наявна	Доступна
2.		Python	Наявна	Доступна
3.		Eviews	Наявна	Доступна
Обрана технологія реалізації ідеї проекту: Python				

4.3 Аналіз ринкових можливостей запуску стартап проекту

Таблиця 4.4 – Попередня характеристика потенційного ринку стартап проекту

№ п/п	Показники стану ринку (найменування)	Характеристика
1.	Кількість головних гравців, од	2
2.	Загальний обсяг продаж, грн/ум.од	7500000

Продовження таблиці 4.4

3.	Динаміка ринку (якісна оцінка)	Зростає
4.	Наявність обмежень для входу (вказати характер обмежень)	Наявність великих гравців у сфері
5.	Специфічні вимоги до стандартизації та сертифікації	Немає
6.	Середня норма рентабельності в галузі (або по ринку), %	30%

Таблиця 4.5 – Характеристика потенційних клієнтів стартап проекту

№ п/п	Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних цільових груп клієнтів	Вимоги споживачів до товару
1.	Потреба в автоматизованій СППР у сфері інвестування	Фізичні особи та невеликі компанії, зацікавлені в інвестуванні	Необхідність невисокої ціни продукту	Вимоги до точності та ефективності роботи СППР
2.		Великі компанії, зацікавлені в інвестуванні	-	

Таблиця 4.6 – Фактори загроз

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
1.	Новий продукт	Потенційні користувачі з підозрою ставляться до нових продуктів	Поширення рекламної кампанії

Таблиця 4.7 – Фактори можливостей

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
1.	Потреба у недорогому продукті	Потреба фізичних осіб та невеликих компаній у недорогій СППР в сфері демографічних процесів	Задоволення потреба фізичних осіб та невеликих компаній у недорогій СППР в сфері демографічних процесів

Таблиця 4.8 – Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
1. Тип конкуренції: олігополія	У сфері домінує невелика кількість компаній	Поширення рекламної кампанії

Продовження таблиці 4.8

2. За рівнем конкурентної боротьби: міжнаціональний	Наявна міжнаціональна конкуренція	-
3. Конкуренція за видами товарів: товарно-видова	Наявна конкуренція між схожими продуктами	-
4. За характером конкурентних переваг: нецінова	Наявна конкуренція завдяки підвищенню якості та надійності продукції	Можливість вийти на ринок з недорогим продуктом
5. За інтенсивністю: не марочна	Наявна конкуренція, де роль торгової марки незначна	-

Таблиця 4.9 – Аналіз конкуренції в галузі за М. Портером

Складові аналізу	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
	СППР SAS СППР Oracle	Доступ до каналів розподілу	-	Розмір закупівель, контроль якості	Ціна
	Висока інтенсивність	Є можливості входу в	Постачальники не	Клієнти диктують	

Продовження таблиці 4.9

Висновки	конкурентної боротьби з боку конкурентів	ринок. Потенційних конкурентів в немає.	диктують умови роботи на ринку.	високі критерії якості продуктів.	-
----------	--	---	---------------------------------	-----------------------------------	---

Таблиця 4.10 – Обґрунтування факторів конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)
1.	Ціна	Ціна запропонованого продукту значно нижче ніж ціни конкуруючих, це значно поширює потенційну клієнтську базу
2.	Ефективність	Результати використання продукту в умовах стохастичної невизначеності є сумірними або кращими за результати конкуруючих продуктів
3.	Поріг входження	Так як у сфері тип конкуренції є олігополією, достатньо складно вивести новий невідомий невеликий продукт на ринок

Таблиця 4.11 – Порівняльний аналіз сильних та слабких сторін проекту

№ п/п	Фактор конкуренто спроможності	Бали 1-20	Рейтинг товарів-конкурентів у порівнянні з запропонованою СППР						
			-3	-2	-1	0	+1	+2	+3
1.	Ціна	15	+						
2.	Ефективність	18			+				
3.	Поріг входження	10							+

Таблиця 4.12 – SWOT аналіз стартап проекту

Сильні сторони: Ціна продукту Ефективність продукту	Слабкі сторони: Невідомість продукту
Можливості: Охоплення аудиторії, що не може дозволити дорогі інтелектуальні комплекси Впровадження інноваційних методів підрахунку ризиків	Загрози: Можлива незацікавленість продуктом через його невідомість та невеликість

Таблиця 4.13 – Альтернативи ринкового впровадження стартап проекту

№ п/п	Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
1.	Розробка програмного забезпечення та грамотна маркетингова програма	Велика	3-5 місяців

4.4 Розроблення ринкової стратегії проекту

Таблиця 4.14 – Вибір цільових груп потенційних клієнтів

№ п/ п	Опис профілю цільової групи потенційни х клієнтів	Готовність споживачів сприйняти продукт	Орієнтовний попит в межах цільової групи (сегменту)	Інтенсивність конкуренції в сегменті	Простота входу у сегмент
1.	Фізичні особи та невеликі компанії, зацікавлені в сфері демографіч них процесів	Значна готовність	Високий	Низька	Середня

Продовження таблиці 4.14

2.	Великі компанії, зацікавлені в сфері демографічних процесів	Незначна готовність	Низький	Висока	Висока
Які цільові групи обрано: Фізичні особи та невеликі компанії, зацікавлені в сфері демографічних процесів					

Таблиця 4.15 – Визначення базової стратегії розвитку

№ п/п	Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентно спроможні позиції відповідно до обраної альтернативи	Базова стратегія розвитку
1.	Розробка програмного забезпечення та грамотна маркетингова програма	За рахунок потреби в дешевому продукті	Ціна	Стратегія спеціалізації

Таблиця 4.16 – Визначення базової стратегії конкурентної поведінки

№ п/п	Чи є проект «першопрхідцем» на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки
1.	Так	Шукати нових споживачів	Буде розроблений продукт зі схожим функціоналом, але іншою методологією	Стратегія заняття конкурентної ніші

Таблиця 4.17 – Визначення стратегії позиціонування

№ п/п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкуренто спроможні позиції власного стартап- проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)
1.	Необхідність невисокої ціни продукту, точності та ефективності роботи СППР	Стратегія спеціалізації	Ціна та ефективність	Низька ціна Висока ефективність Простота у використанні

4.5 Розроблення маркетингової програми стартап проекту

Таблиця 4.18 – Визначення ключових переваг потенційного товару

№ п/п	Потреба	Вигода товару	Ключові переваги перед конкурентами
1.	Потреба в недорогому продукті	Пропонує недорогий продукт	Ціна товару нижче ніж ціна товару конкурентів
2.	Потреба в ефективному продукті	Пропонує ефективний продукт	Ефективність товару вище ніж ефективність товару конкурентів

Таблиця 4.19 – Опис трьох рівнів моделі товару

Рівні товару	Сутність та складові		
I. Товар за задумом	Недорога СППР в сфері демографічних процесів в умовах стохастичної невизначеності		
II. Товар у реальному виконанні	Властивості/характеристики	М/Нм	Вр/Тх /Тл/Е/Ор
	1. Мульти платформенність 2. Зручний інтуїтивний інтерфейс	-	-
	Якість: стандарти ефективності		
	Пакування: електронне розповсюдження		
	Марка: IDAV InvestStoch		

Продовження таблиці 4.19

III. Товар із підкріпленням	До продажу
	Після продажу технічна підтримка
За рахунок чого потенційний товар буде захищено від копіювання: захист інтелектуальної власності	

Таблиця 4.20 – Визначення меж встановлення ціни

№ п/п	Рівень цін на товари-замінники	Рівень цін на товари - аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
1.	-	7500000-54000000 грн.	25000+ грн.	8000-18000 грн.

Таблиця 4.21 – Формування системи збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
1.	Закупівля через інтернет	Підтримка нормального функціонування сайту	0	Електронне розповсюдження

Таблиця 4.22 – Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комуніка цій цільових клієнтів	Ключові позиції, обрані для позиціонуван ня	Завдання рекламного повідомлення	Концепція рекламного звернення
1.	Потреба в недорогому та ефективному продукті	Інтернет- мережі	Низька ціна Висока ефективність Простота у використанні	Провести якісну маркетингову кампанію	Донести специфіку продукту

Висновки

Є можливість ринкової комерціалізації проекту (наявний попит, наявна динаміка ринку, наявна рентабельність роботи на ринку). Є перспективи впровадження з огляду на потенційні групи клієнтів, бар'єри входження, стан конкуренції, конкурентоспроможність проекту.

Доцільно обрати альтернативу розробки програмного забезпечення та грамотної маркетингової програми для ринкової реалізації проекту. Подальша імплементація проекту є доцільною.

ВИСНОВКИ

Кожного дня наша планета поповнюється новими жителями і, на жаль, також несе втрати, відбуваються міграції з одного місця в інше, відбуваються явища, що змінюють світ та наш звичний образ життя. Саме ці динамічні процеси творять наше повсякдення. Вміння проаналізувати та спрогнозувати їх поведінку дає величезні можливості для побудови стратегій ведення внутрішньої та зовнішньої політики держав, для того щоб покращити рівень життя, та підняти у рейтингу на світовій арені.

У даній роботі було проаналізовано демографічні процеси нашої країни. Досліджено ряд моделей прогнозування нелінійних нестационарних процесів, моделей апроксимації прогнозування та можливість використання нейронних мереж. Висвітлено переваги та недоліки вже існуючих популярних пакетних рішень створення прогнозу.

Було сформовано методику побудови моделей інтегрованих та гетероскедастичних процесів, обрано найкращу модель для прогнозування демографічних процесів та найголовніше – реалізовано прогноз демографічного стану України, та проведено аналіз отриманих результатів. На основі вище переліченого, було висвітлено основні ознаки, що впливають на демографічний розвиток України.

ПЕРЕЛІК ПОСИЛАНЬ

1. Бідюк П.І., Половцев О.В. Аналіз та моделювання економічних процесів перехідного періоду. К: ПЛАБ -75, 1999. 230 с.
2. Бідюк П.І. Системний підхід до побудови математичних моделей на основі часових рядів. *Системні дослідження та інформаційні технології*, №3, 2002. 114-131 с.
3. Бідюк П.І. Часові ряди: моделювання та прогнозування. Київ: ЕКМО, 2004. 144 с.
4. Шекера О.Г. Демографічна ситуація у світі та в Україні. Науково-практичний журнал «Здоров'я суспільства», 2014. 123 с.
5. Капіца С. П. Общая теория роста населения Земли. М.: ММВБ, 2009. 120 с.
6. Власенко Н.С., Макарова О.В., Пирожков С.І., та інші. Комплексний демографічний прогноз України на період до 2050 р. за ред. член-кореспондент НАНУ, д.е.н., проф. Е.М. Лібанової. К.: Український центр соціальних реформ, 2006. 138 с.
7. Цвігун І.А. Демографічна безпека України та напрями її регулювання: монографія. Кам'янець-Подільський: Видавець ПП Зволейко Д.Г., 2013. 400 с.
8. Пальян З. О. Навчальний посібник Демографічна статистика. Київський Національний Економічний Університет України, 2003. 167 с.
9. Капица С. П. Феноменологическая теория роста населения Земли. М.: Наука, 1996. 80 с.
10. Населення України, 2015 рік : демогр. Щорічник. Держ. ком. статистики України, Упр. статистики населення; Л. М. Стельмах (відп. за вип.). К.: Консультант, 2015. 466 с.

11. Борисов В. А., Синельников А.Б. Демографический анализ М.: НИИ семьи, 1996. 66 с.
12. Полянська-Василенко Н. В. Історія України. К.: Либідь, 1995. 588 с.
13. Розподіл постійного населення України за статтю та віком на 1 січня 2019 року: Державна служба статистики. К., 2019. URL: http://database.ukrcensus.gov.ua/PXWEB2007/ukr/news/op_popul.asp (дата звернення: 09.09.2020)
14. Міграційний рух населення: Державна служба статистики. К., 2019. URL: http://database.ukrcensus.gov.ua/PXWEB2007/ukr/news/op_migr.asp (дата звернення 09.09.2020)
15. Формування приросту (скорочення) чисельності наявного населення та показники природного руху населення в Україні. К., 2019 URL: http://database.ukrcensus.gov.ua/PXWEB2007/ukr/news/op_n_mov.asp (дата звернення 10.09.2020).
16. П. І. Бідюк, О. І. Савенков, І. В. Баклан Часові ряди: моделювання і прогнозування. К.: ЕКМО, 2003. 144 с.
17. Андерсон Т. Статистический анализ временных рядов. М.: Мир, 1976. 755 с.
18. Лук'яненко Г. Г. Сучасні економетричні методи у фінансах. Навчальний посібник. К.: Літера ЛТД, 2002. 352 с.
19. Lindley D.V. Making Decisions. New York: Wiley, 2004. 540 p.
20. П. І. Бідюк, І. В. Баклан Системні дослідження та інформаційні технології. К.: ЕКМО 2002. 131 с.
21. Challis, R. E., and Kitney, R. I. "Biomedical signal processing (in four parts). Part 1. Time-domain methods. *Medical & Biological Engineering & Computing*. 1991. 509-524 p.

22. Лукашин Ю.П. Адаптивные методы долгосрочного прогнозирования. М.: Финансы и статистика, 2003. 414 с.
23. Зельнер А. Методы в эконометрии. Москва: Статистика, 1980. 438 с.
24. Chatfield C. Time series forecasting. London: Chapman & Hall, 2000. 267 p.
25. Згуровский М.З., Подладчиков В.Н. Аналитические методы калмановской фильтрации. Київ: Наукова думка, 1995. 285 с.
26. Fryzlewicz, P., Van Bellegem, S. & von Sachs, R. Forecasting non-stationary time series by wavelet process modelling. *Annals of the Institute of Statistical Mathematics*. 2003. 737-764 p. URL: <https://link.springer.com/article/10.1007/BF02523391> (Дата звернення 12.10.2020).

ДОДАТОК А. Динаміка чисельності населення

Таблиця А.1 – Динаміка чисельності населення [15].

	Кількість наявного населення		
	всього, тис.осіб	у тому числі	
		Міське	Сільське
1991	51 944 400	35 085 200	16 859 200
1992	52 056 600	35 296 900	16 759 700
1993	52 244 100	35 471 000	16 773 100
1994	52 114 400	35 400 700	16 713 700
1995	51 728 400	35 118 800	16 609 600
1996	51 297 100	34 767 900	16 529 200
1997	50 818 400	34 387 500	16 430 900
1998	50 370 800	34 048 200	16 322 600
1999	49 918 100	33 702 100	16 216 000
2000	49 429 800	33 338 600	16 091 200
2001	48 923 200	32 951 700	15 971 500
2002	48 457 102	32 574 371	15 882 731
2003	48 003 463	32 334 120	15 669 343
2004	47 622 434	32 148 345	15 474 089
2005	47 280 817	32 009 320	15 271 497
2006	46 929 525	31 877 710	15 051 815
2007	46 646 046	31 777 367	14 868 679
2008	46 372 664	31 668 757	14 703 907
2009	46 143 714	31 587 203	14 556 511
2010	45 962 947	31 524 795	14 438 152
2011	45 778 534	31 441 649	14 336 885
2012	45 633 637	31 380 874	14 252 763
2013	45 553 047	31 378 639	14 174 408
2014	45 426 249	31 336 623	14 089 626

Продовження таблиці А.1

2015	42 929 298	29 673 113	13 256 185
2016	42 760 516	29 584 952	13 175 564
2017	42 584 542	29 482 313	13 102 229
2018	42 356 786	29 377 253	13 002 134
2019	42 153 335	29 276 378	12 934 224
2020	42 002 123	29 123 432	12 884 213

ДОДАТОК Б. Динаміка демографічного розвитку населення

Таблиця Б.1 – Динаміка демографічного розвитку населення [15].

	Кількість народжених	Частка дітей народжених жінками, які не перебували у зареєстрованому шлюбі, %	Кількість померлих	Природний приріст населення	Міграція населення між Україною та іншими державами приріст населення, тис. Одиниць	Кількість зареєстрованих шлюбів, одиниць	Кількість зареєстрованих розлучень, одиниць
1995	492 861	13,2	792 587	-299 726	-131,6	431 731	198 300
1996	467 211	13,6	776 717	-309 506	-169,2	307 543	193 030
1997	442 581	15,2	754 151	-311 570	-136,0	345 013	188 232
1998	419 238	16,2	719 954	-300 716	-152,0	310 504	179 688
1999	389 208	17,4	739 170	-349 962	-138,3	344 888	175 781
2000	385 126	17,3	758 082	-372 956	-33,8	274 523	197 274
2001	408 589	18,0	745 952	-369 474	-24,4	309 602	181 334
2002	427 259	19,0	754 911	-364 223	-7,6	317 228	183 538
2003	426 086	19,9	765 408	-356 819	4,6	370 966	177 183
2004	460 368	20,4	761 261	-334 002	14,2	278 225	173 163
2005	472 657	21,4	781 961	-355 875	16,8	332 143	183 455
2006	510 589	21,1	758 092	-297 724	14,9	354 959	179 123
2007	512 525	21,4	762 877	-290 220	13,4	416 427	178 364
2008	497 689	20,9	754 460	-243 871	16,1	321 992	166 845

Продовження таблиці Б.1

2009	502 595	21,2	706 739	–194 214	17,1	318 198	145 439
2010	520 705	21,9	698 235	–200 546	61,8	305 933	126 068
2011	503 657	21,9	664 588	–161 993	31,9	355 880	182 490
2012	465 882	21,4	663 139	–142 434	21,1	278 276	168 508
2013	411 781	22,1	662 368	–158 711	9,3	304 232	164 939
2014	408 589	21,1	632 296	–166 414	–33,8	294 962	130 673
2015	427 259	20,6	594 796	–183 015	–24,2	299 038	129 373
2016	397 037	22,1	583 631	–186 594	–19,7	229 453	129 997
2017	363 987	19,7	574 123	–210 136	–20,1	249 522	128 734

ДОДАТОК В. Середня тривалість життя жінок і чоловіків

Таблиця В.1 – Середня тривалість життя жінок і чоловіків [14].

Період, за який розраховані показники (роки)	Середня очікувана тривалість життя при народженні обидві статі, років	у тому числі	
		Чоловіки	Жінки
1991	69,56	64,62	74,21
1992	68,97	63,81	73,98
1993	68,29	63,16	73,35
1994	67,66	62,39	72,95
1995	66,79	61,22	72,54
1996	67,08	61,52	72,80
1997	67,66	62,23	73,19
1998	68,50	63,17	73,84
1999	68,07	62,62	73,61
2000	67,72	62,10	73,53
2001	67,89	62,32	73,63
2002	68,32	62,70	74,13
2003	68,24	62,64	74,06
2004	68,22	62,60	74,05
2005	67,96	62,23	73,97
2006	68,10	62,38	74,06
2007	68,25	62,51	74,22
2008	68,27	62,51	74,28
2009	69,29	63,79	74,86
2010	70,44	65,28	75,50
2011	71,02	65,98	75,88
2012	71,15	66,11	76,02

Продовження таблиці В.1

2013	71,37	66,34	76,22
2014	71,37	66,25	76,37
2015	71,38	66,37	76,25
2016	71,48	66,49	76,27
2017	71,43	66,52	76,38

ДОДАТОК Г. Лістинг програми

```

import pandas as pd # to read data
import numpy as np # for calculations and data manipulations
from sklearn.preprocessing import PolynomialFeatures # for data manipulations
from sklearn.neural_network import MLPRegressor # for studying neural network

# reading data
data = pd.read_csv("data.csv")[["year", "population"]].astype("float")
print("Data is read", end="\n\n")

# manipulations with data
X = pd.DataFrame(index=data.index).astype("float")
X["year"] = data["year"]
X["cbrt"] = np.cbrt(data["year"])
X["sqrt"] = np.sqrt(data["year"])
X["log"] = np.log(data["year"])
X = PolynomialFeatures(degree=2).fit_transform(X)
y = data["population"]

# studying of network
print("Studying of neural network")
NW = MLPRegressor((125, 125, 125, 125, 125, 125), activation="relu").fit(X, y)
print("Network is studied", end="\n\n")

# prediction of population
print("Prediction")
pred = []
for i in [2021, 2022, 2023, 2024, 2025]:
    pred.append([i, np.cbrt(i), np.sqrt(i), np.log(i)])
pred = PolynomialFeatures(degree=2).fit_transform(pred)
prediction = NW.predict(pred)

# building and show table with predicted data
minPopularity = [ abs(int(num*10))-y.std()/len(y) for num in prediction ]
maxPopularity = [ abs(int(num*10))+y.std()/len(y) for num in prediction ]
prediction = pd.DataFrame(index=[2021, 2022, 2023, 2024, 2025])
prediction["minPopularity"] = minPopularity
prediction["maxPopularity"] = maxPopularity
print(prediction)

```

`input()`